Qubit-channel metrology with very noisy initial states

David Collins[*]

Department of Physical and Environmental Sciences, Colorado Mesa University, Grand Junction, Colorado 81501, USA

We consider an arbitrary qubit channel depending on a single parameter, which is to be estimated by a physical process. Using the quantum Fisher information per channel invocation to quantify the estimation accuracy, we consider various estimation protocols when the available initial states are mixed with very low purity r. We compare a protocol using a single channel invocation on one out of n qubits prepared in a particular correlated input state to the optimal protocol using uncorrelated input states, with the same initial-state purity. We show that, to lowest order in initial-state purity, for a unital channel this correlated-state protocol enhances the estimation accuracy by a factor between $n - 1$ and n, provided that $nr^2 \ll 1$. We also show that to lowest order in initial-state purity, a broad class of nonunital channels yields no gain regardless of the input state.

I. INTRODUCTION

Quantum parameter estimation, or metrology, considers using physical quantum systems as measuring devices. Typically a system is prepared in a known state and is then subjected to an evolution of a known type but which depends on an unknown parameter that is to be estimated. The parameter must subsequently be inferred from measurements on the system. Classical statistics and quantum physics constrain the success of such procedures; combining these has led to a quantum estimation framework [1–9].

This has been applied to various situations, including estimation of parameters in phase-shifts channels [5], depolarizing channels [10–12], Pauli channels [13,14], and amplitude damping channels [15]. A key issue is whether using states only available to quantum systems (such as entangled states) enhances the estimation accuracy compared to that for "classical" repeat and average strategies using uncorrelated states. Sometimes this is true.

Most studies focus on the absolute optimal situations, which require pure initial states. However, in some situations such as solution-state nuclear magnetic resonance (NMR), pure states are unavailable and the issue becomes whether advantages arise when correlating mixed or noisy states that would otherwise be used in an uncorrelated estimation protocol. This has been addressed for the phase-shift [16–19], phase-flip [20], and depolarizing channels [21]. These studies considered the situation where all available qubits are initially in mixed states and focused on the enhancement of estimation accuracy in terms of the quantum Fisher information. So far there is no general result for all single-qubit channels and for all initial state purities; each study yielded a distinct algebraic expression that appears difficult to interpret for a general degree of purity in the initial states. However it was observed that this simplifies considerably for the qubit phase-flip and depolarizing channels [20,21] when the initial-state purity is

very low and, for a particular protocol which uses correlated states, the accuracy can be enhanced by a factor equal to the number of qubits available.

The question that this article addresses is whether, for any single-qubit channel, there exists a parameter estimation protocol that uses correlated states and provides such accuracy enhancements when the purity of the available initial states is extremely low. While this is partly motivated by the answers from previous studies of specific channels, a physical motivation comes from NMR, where the available purity of the states is very low. A generic model for this could be an ensemble of identical molecules each with the same number of spin-1/2 nuclei and a channel acting on the same nuclear spin in each molecule. Would there be a parameter estimation advantage to correlating the spins within each molecule prior to the channel action?

We note that the situation where some qubits are initially in pure states and others are in mixed states has been studied and also shows advantages for estimation [22]; we do not consider this case here.

This article is organized as follows. Section II reviews the general description of qubit system evolution and describes the basic notions of parameter estimation. Section III reviews quantum estimation theory and adapts this to situations where the available initial states are very noisy. Section IV applies this to a single-qubit protocol, which serves as a baseline for comparison with a multiple qubit correlated-state protocol that is described in Sec. V. This contains the main results of this article; these are summarized in Sec. VI.

II. SINGLE PARAMETER QUBIT-CHANNEL METROLOGY

We consider general single-qubit quantum channels. Prior to evolution, the *channel input* state for a single qubit can be represented as

$$\hat{\rho}_i = \tfrac{1}{2}\big(\hat{I} + r\hat{\sigma}_{\mathbf{r}_i}\big), \tag{1}$$

———
[*]dacollin@coloradomesa.edu

where $\hat{I}$ is the identity operator, $\mathbf{r}_i$ is the *input state Bloch-sphere direction,* a three-dimensional real unit vector, and $\hat{\sigma}_{\mathbf{r}_i} = \mathbf{r}_i \cdot \hat{\boldsymbol{\sigma}} = r_{ix}\hat{\sigma}_x + r_{iy}\hat{\sigma}_y + r_{iz}\hat{\sigma}_z$ (throughout this article we use the notation $\hat{\sigma}_{\mathbf{a}} := \mathbf{a} \cdot \hat{\boldsymbol{\sigma}}$, where $\mathbf{a}$ is any real three-dimensional vector). Here r, which satisfies $0 \leqslant r \leqslant 1$, is called the purity of the state and quantifies the mixedness or noisiness of the state. Under the channel $\hat{\rho}_i \mapsto \hat{\rho}_f$ and, again generically, $\hat{\rho}_f = (\hat{I} + \hat{\sigma}_{\mathbf{r}_f})/2$, where $\mathbf{r}_f$ is the final state Bloch-sphere direction with $|\mathbf{r}_f| \leqslant 1$. For any channel [23],

$$\mathbf{r}_f = M(r\mathbf{r}_i) + \mathbf{d} = rM\mathbf{r}_i + \mathbf{d}, \qquad (2)$$

where M is a 3×3 real *Bloch-sphere matrix* $\mathbf{d}$ and a real *Bloch-sphere shift vector*. In Appendix A we show that $|\mathbf{d}| \leqslant 1$ and $|\mathbf{d}| = 1$ is only possible when $M = 0$.

By linearity, the channel maps $\hat{I} \mapsto \hat{I} + \hat{\sigma}_{\mathbf{d}}$ and also $\hat{\sigma}_{\mathbf{r}_i} \mapsto M\mathbf{r}_i \cdot \hat{\boldsymbol{\sigma}}$.

Qubit channels for which $\mathbf{d} = 0$ are called unital; these map $\hat{I} \mapsto \hat{I}$. Examples are unitary channels, Pauli channels, and the depolarizing channel. Nonunital channels, for which $\mathbf{d} \neq 0$, include the amplitude damping channel.

We consider channels which depend, via only M and $\mathbf{d}$, on a single parameter λ, which is independent of the channel input state. The task will be to estimate the parameter by a physical process in which one or more qubits, prepared in known input states, undergo evolution via one or more identical copies of the channel. The channel actions are followed by measurements, whose outcomes are used to infer the parameter. The goal will be to choose input states, measurements, and statistical inference processes that minimize fluctuations in the estimates they generate; we assume that the key cost of such procedures is the number of channel invocations.

III. ENTANGLEMENT-ASSISTED METROLOGY WITH NOISY INITIAL STATES

A standard formalism for assessing physical quantum estimation uses the density operator $\hat{\rho}_f(\lambda)$ for the (possibly multiple qubit) entire system immediately after the final channel invocation. The estimate λ_{est} is inferred from measurement outcomes via a known estimator function. We will require that this estimator is unbiased, i.e., the mean of the estimates equals the true parameter value. Then the classical Cramér-Rao bound (CRB) bounds the variance in the estimate via $\text{var}\,(\lambda_{est}) := \langle (\lambda_{est} - \langle \lambda_{est} \rangle)^2 \rangle \geqslant 1/F(\lambda)$ regardless of the choice of estimator [6,24] (the angle brackets indicating the mean over all possible measurement outcomes). The classical Fisher information

$$F(\lambda) := \int \left[\frac{\partial \ln p(x_1, x_2 \dots |\lambda)}{\partial \lambda} \right]^2 p(x_1, x_2 \dots |\lambda)\, dx_1 dx_2 \dots,$$
$$(3)$$

is determined from the probability distribution $p(x_1, x_2, \dots, |\lambda)$ for the process measurement outcomes, $x_1, x_2, \dots$. There is always an estimator which asymptotically attains the lower bound [24].

In quantum estimation the choice of measurement affects the probability distribution used to compute the classical Fisher information. However, a further constraint is given by the quantum Cramér-Rao bound (QCRB) $F(\lambda) \leqslant H(\lambda)$,

where the quantum Fisher information (QFI) is

$$H(\lambda) = \text{Tr}\,[\hat{\rho}_f(\lambda)\hat{L}^2(\lambda)], \qquad (4)$$

and $\hat{L}(\lambda)$ is the symmetric logarithmic derivative (SLD) defined via

$$\frac{\partial \hat{\rho}_f(\lambda)}{\partial \lambda} = \frac{1}{2}\,[\hat{L}(\lambda)\hat{\rho}_f(\lambda) + \hat{\rho}_f(\lambda)\hat{L}(\lambda)]. \qquad (5)$$

The SLD and the QFI only depend on the premeasurement system state and thus $\text{var}\,(\lambda_{est}) \geqslant 1/H(\lambda)$ offers a bound that is independent of both the choice of measurement and estimator [3,6,25–27].

The SLD can always be computed from a diagonal decomposition $\hat{\rho}_f(\lambda) = \sum_j p_j(\lambda)|\phi_j(\lambda)\rangle\langle\phi_j(\lambda)|$, via [6,9]

$$\hat{L}(\lambda) = 2 \sum_{j,k} \frac{\langle \phi_j|\dot{\hat{\rho}}_f|\phi_k\rangle}{p_j + p_k}\,|\phi_j\rangle\langle\phi_j|, \qquad (6)$$

where the dot indicates differentiation with respect to the parameter. In some cases there are simpler algebraic methods for computing the SLD [20]. Also, simple matrix algebra and Eqs. (4) and (5) give

$$H(\lambda) = \text{Tr}\left[\frac{\partial \hat{\rho}_f(\lambda)}{\partial \lambda}\hat{L}(\lambda) \right]. \qquad (7)$$

It is always possible to saturate the QCRB by choosing a projective measurement in the eigenbasis of the SLD but it cannot be assured that this choice is independent of the unknown parameter [6,26,28]. In such cases, there exist various other measurement schemes that asymptotically saturate the QCRB [29].

Thus the QFI quantifies the accuracy of possible physical measurement procedures. Generally the task in any quantum parameter estimation study has been to engineer a final system state that maximizes the QFI, subject to various system constraints and resources (i.e., number of channel invocations, number of available systems, types of initial states available, etc.) [3–6,8–10,12–15,17,18,20,21,25,30–34]. We will adopt the common approach in which the only costly resource is the number of channel invocations. We therefore aim to evaluate estimation protocols via the QFI per channel invocation.

A protocol that uses multiple copies of the channel might enhance the QFI. Indeed, an *independent channel invocation protocol* which invokes the channel once on each of the m systems prepared independently in the same input state, as illustrated in Fig. 1(a), gives exactly an m-fold increase in the QFI [25]. Here the QFI per channel invocation is the same as a procedure in which the channel is invoked once on a single system.

In contrast, entanglement-assisted metrology considers protocols, illustrated in Fig. 1(b), where the available quantum systems are prepared in an entangled or otherwise correlated state and thereafter the channel is invoked on a subset of these while the remaining ancillary systems function as spectators in a noiseless environment. Such entanglement assistance can enhance the QFI per channel invocation (versus uncorrelated or independent protocols) [9,34–37]. A key issue in quantum metrology is to establish when and to what extent entanglement assisted protocols can assist parameter estimation.
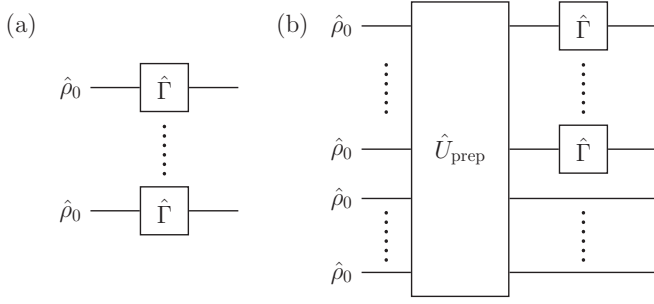
FIG. 1. Two quantum metrology protocols. (a) A classical independent channel invocation protocol in which the channel $\hat{\Gamma}(\lambda)$ is invoked on m systems each in the same input state $\hat{\rho}_0$. (b) An entanglement-assisted protocol using n systems, prepared in an entangled or correlated state via a multisystem preparatory unitary $\hat{U}_{\text{prep}}$ and with the channel invoked once on each of m of these. The lower $n - m$ systems serve as ancillas.

We address this for parameter estimation for single-qubit channels subject to the following considerations. First, we assume that a fixed finite number of qubits n is available and each is initially in the same *initial state*; generically this can be expressed as $\hat{\rho}_0 = (\hat{I} + r\hat{\sigma}_{\mathbf{r}_0})/2$, where $\mathbf{r}_0$ is the initial-state Bloch-sphere direction unit vector and r is the purity. Second, we will assume that the purity is very small. Specifically, as will be shown later, our analysis is valid when $r \ll 1/\sqrt{n}$. Third, we assume that a single parameter-independent preparatory unitary $\hat{U}_{\text{prep}}$ is applied to the entire system of qubits. This produces a channel *input state* $\hat{\rho}_{\text{i}} := \hat{U}_{\text{prep}}\hat{\rho}_0^{\otimes n}\hat{U}_{\text{prep}}^{\dagger}$, where $\hat{\rho}_0^{\otimes n}$ is the initial state for the system of all n qubits. Fourth, we assume that after the preparatory unitary the channel $\hat{\Gamma}(\lambda)$ is invoked once on a single qubit, mapping the channel input state to a final premeasurement state via $\hat{\rho}_{\text{i}} \overset{\hat{\Gamma}(\lambda)}{\mapsto} \hat{\rho}_{\text{f}}(\lambda)$. The entire process is illustrated in Fig. 2(b). We term this a *correlated-state* protocol since for low enough purity the states may be separable and the presence of entanglement is not assured (for details of separability in previously studied specific cases see [20,21]).

We will compare such a correlated-state protocol to one in which there is only a single qubit available and the channel is invoked once, as illustrated in Fig. 2(a); we call the the latter the single-qubit, single-channel (SQSC) protocol. This yields the same QFI per channel invocation as an independent-channel protocol and thus in terms of the QFI per channel invocation, we are effectively comparing a correlated-state protocol to an independent-channel protocol. Therefore we

ask whether, for given initial-state purity, there is an entangling preparatory unitary so that the correlated-state protocol of Fig. 2(b) yields a QFI exceeding than that of any SQSC protocol, for the same initial-state purity, and, to lowest nontrivial order in the initial-state purity, what gains such an correlated-state protocol provides.

One physical motivation for this would be estimation of parameters associated with the evolution of single spins in room-temperature, solution-state NMR. Each molecule contains n distinct nuclear spins; intermolecular interactions average to zero and can be ignored. The entire ensemble only serves to amplify measurement signals and provide representative sampling of measurement outcomes. Therefore preparatory unitaries, channel actions, and measurements may be regarded as restricted to within individual molecules. Typically [23] $r \approx 10^{-4}$ and thus our analysis applies whenever the number of nuclear spins within one molecule would be much less than 10^8. Most current room-temperature solution-state NMR lies well within this realm.

Ultimately it may be of interest to compare correlated-state protocols in which the channel is invoked more than once. However, it is known that for the particular cases of the phase-flip [20] and depolarizing channels [21] such correlated-state protocols do not yield advantages for all parameter values when there are two or more channel invocations. On the other hand, when there is only one channel invocation, these protocols definitely are advantageous over all parameter values for the phase-flip channel [20] and are probably so for the depolarizing channel [21]. Thus, we only consider the situation where the channel is invoked once and assess whether the remaining ancillary qubits assist in the parameter estimation.

Sometimes low purity situations can be assessed using expressions for the QFI that are valid for all purities [18,20,21]. Such exact expressions do not appear to be available for all channels and protocols. We will present a series approach for computing the QFI that is correct to the lowest nontrivial orders in the purity.

Before doing this, we consider what the framework by Escher *et al.* [32] for noisy quantum metrology describes for this situation. The Escher framework considers a system in an initial *pure* state $\hat{\rho}_0$. The channel plus noise acts on this to produce a final premeasurement state, which can be expressed [23] using a set of Kraus operators $\{\hat{\Pi}_l\}$ as $\hat{\rho}_{\text{f}}(\lambda) = \sum_l \hat{\Pi}_l(\lambda)\hat{\rho}_0\hat{\Pi}_l^{\dagger}(\lambda)$. Then the QFI is bounded [32] according to

$$H \leqslant 4\{\text{Tr}\,(\hat{H}_1\hat{\rho}_0) - [\text{Tr}\,(\hat{H}_2\hat{\rho}_0)]^2\}, \tag{8}$$

where $\hat{H}_1 := \sum_l \frac{\partial \hat{\Pi}_l^{\dagger}}{\partial \lambda}\frac{\partial \hat{\Pi}_l}{\partial \lambda}$ and $\hat{H}_2 := i\sum_l \frac{\partial \hat{\Pi}_l^{\dagger}}{\partial \lambda}\hat{\Pi}_l$. For any given channel plus noise there are multiple Kraus representations and that which minimizes the right-hand side of Eq. (8) exactly yields the QFI [32]. The Escher framework does not give an explicit method for finding the optimal Kraus representation. Sometimes careful choices of Kraus operators yield useful bounds on the QFI [32]. This framework can address single-qubit channel parameter estimation with noisy initial states by regarding the input state for each qubit as produced from a pure state via a depolarizing channel. This gives Kraus operators that combine the channel and initial depolarization Kraus operators and these can be used to compute the
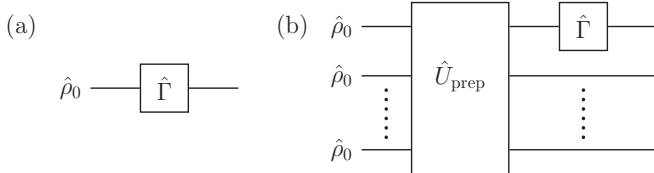

FIG. 2. Single channel qubit metrology protocols. (a) The SQSC protocol with a single channel invocation on a single qubit. (b) An entanglement-assisted protocol using n qubits with the channel invoked once on one of them.

right-hand side of Eq. (8). As shown in Appendix B, applying this to the example of a phase-flip channel acting on a single qubit gives a bound that is in excess of the known QFI for this channel acting on a noisy input state; in fact the bound does not even refer to the initial-state purity. This is a consequence of the choice of Kraus representation and there must clearly be a more restrictive choice, but what that is is not apparent and therefore the Escher framework is not immediately instructive in cases where the noise appears in initial states.

A. Series computation of the QFI

If the initial-state purity is sufficiently low then we can approximate the QFI to lowest order in the purity by expressing the QFI and its constituent ingredients as series in increasing powers of the purity; this circumvents difficulties associated with exact computation of the SLD.

To do so, the initial state can be expressed as

$$\hat{\rho}_0^{\otimes n} = \sum_{j=0}^{n} r^j \, \hat{\rho}_0^{(j)}, \tag{9}$$

where $\hat{\rho}_0^{(j)}$ is an operator independent of the purity. The preparatory unitary maps this to the input state $\hat{\rho}_i := \hat{U}_{\text{prep}} \hat{\rho}_0^{\otimes n} \hat{U}_{\text{prep}}^\dagger$, and thus

$$\hat{\rho}_i = \sum_{j=0}^{n} r^j \, \hat{\rho}_i^{(j)}, \tag{10}$$

where $\hat{\rho}_i^{(j)} = \hat{U}_{\text{prep}}^\dagger \hat{\rho}_0^{(j)} \hat{U}_{\text{prep}}$ is again independent of r. Similarly the final state can be expressed as

$$\hat{\rho}_f(\lambda) = \sum_{j=0}^{n} r^j \, \hat{\rho}_f^{(j)}(\lambda), \tag{11}$$

where $\hat{\rho}_f^{(j)}(\lambda)$ is completely determined by evaluating the channel actions on $\hat{\rho}_i^{(j)}$.

Given that the state for each qubit is $\hat{\rho}_0 = (\hat{I} + r\hat{\sigma}_{\mathbf{r}_0})/2$, two lowest order initial-state terms for the n qubit system are

$$\hat{\rho}_0^{(0)} = \frac{1}{N} \, \hat{I}^{\otimes n}, \tag{12a}$$

$$\hat{\rho}_0^{(1)} = \frac{1}{N} \left[\hat{\sigma}_{\mathbf{r}_0} \otimes \hat{I}^{\otimes(n-1)} + \hat{I} \otimes \hat{\sigma}_{\mathbf{r}_0} \otimes \hat{I}^{\otimes(n-2)} \right.$$
$$\left. + \cdots + \hat{I}^{\otimes(n-1)} \otimes \hat{\sigma}_{\mathbf{r}_0} \right], \tag{12b}$$

where $N = 2^n$. For any preparatory unitary $\hat{\rho}_i^{(0)} = \hat{I}^{\otimes n}/N$ since $\hat{U}_{\text{prep}} \hat{I} \hat{U}_{\text{prep}}^\dagger = \hat{I}$. Higher order terms in $\hat{\rho}_i$ depend on the preparatory unitary. There are no simple general expressions for the lowest order terms in the final state as certain channels, such as the amplitude damping channel map the identity in a nontrivial way.

Similarly the SLD and QFI can be expressed as power series, possibly with infinitely many terms, in r. Thus

$$\hat{L}(\lambda) = \sum_{j=0}^{\infty} r^j \, \hat{L}^{(j)}(\lambda), \tag{13}$$

where $\hat{L}^{(j)}(\lambda)$ is an operator independent of r, and

$$H = \sum_{j=0}^{\infty} r^j \, H^{(j)}, \tag{14}$$

where $H^{(j)}$ is independent of r. The operators $\hat{L}^{(j)}(\lambda)$ can be evaluated by substituting from Eqs. (11) and (13) into Eq. (5). The result must be true for all values of r and comparing terms order by order gives

$$\frac{\partial \hat{\rho}_f^{(k)}}{\partial \lambda} = \frac{1}{2} \sum_{j=0}^{k} \left(\hat{L}^{(k-j)} \hat{\rho}_f^{(j)} + \hat{\rho}_f^{(j)} \hat{L}^{(k-j)} \right). \tag{15}$$

Similarly substituting from Eqs. (11) and (13) into Eq. (7) gives

$$H^{(j)} = \sum_{k=0}^{j} \text{Tr} \left[\frac{\partial \hat{\rho}_f^{(j-k)}}{\partial \lambda} \hat{L}^{(k)} \right]. \tag{16}$$

This allows for an iterative calculation of the QFI in increasing orders of the purity parameter; for sufficiently low purities, the QFI can be approximated by truncation.

It is also useful to determine series expressions for the eigenstates of the SLD in order to assess possible measurements that saturate the quantum CRB. Denote the normalized eigenstate of the SLD by $|\phi\rangle$ and the associated eigenvalue by μ. Again these can be expanded as power series in r, giving

$$|\phi\rangle = \sum_{j=0}^{\infty} r^j |\phi^{(j)}\rangle \tag{17}$$

and

$$\mu = \sum_{j=0}^{\infty} r^j \mu^{(j)}. \tag{18}$$

The normalization condition $\langle \phi | \phi \rangle = 1$ must hold for all r and implies that $\langle \phi^{(0)} | \phi^{(0)} \rangle = 1$. Then order-by-order comparison of terms in $\hat{L}|\phi\rangle = \mu|\phi\rangle$ gives, for each $k = 0, 1, \ldots$,

$$\sum_{j=0}^{k} \hat{L}^{(k-j)} |\phi^{(j)}\rangle = \sum_{j=0}^{k} \mu^{(k-j)} |\phi^{(j)}\rangle. \tag{19}$$

This yields an iterative scheme for determining the eigenstates of the SLD and hence one possible projective measurement which saturates the quantum CRB.

B. Unital channels

A unital channel maps $\hat{I} \xrightarrow{\Gamma} \hat{I}$ and here $\hat{\rho}_f^{(0)} = \hat{I}^{\otimes n}/N$. Repeatedly using Eq. (15) results in

$$\hat{L}^{(0)} = 0, \tag{20a}$$

$$\hat{L}^{(1)} = N \frac{\partial \hat{\rho}_f^{(1)}}{\partial \lambda}. \tag{20b}$$

Thus, for unital channels, Eq. (16) yields

$$H^{(0)} = 0, \tag{21a}$$

$$H^{(1)} = 0, \tag{21b}$$

$$H^{(2)} = N \text{Tr} \left[\left(\frac{\partial \hat{\rho}_f^{(1)}}{\partial \lambda} \right)^2 \right]. \tag{21c}$$

This immediately establishes the result, found earlier [18,20,21] for the qubit phase-shift, phase-flip, and depolarizing channels, that the lowest order terms for the QFI are second order in the purity, provided that the preparation step consists of unitary operations only.

Additionally Eq. (19) allows for computation of the eigenstates of the SLD via

$$\hat{L}^{(0)}|\phi^{(0)}\rangle = \mu^{(0)}|\phi^{(0)}\rangle, \tag{22a}$$

$$\hat{L}^{(1)}|\phi^{(0)}\rangle + \hat{L}^{(0)}|\phi^{(1)}\rangle = \mu^{(0)}|\phi^{(1)}\rangle + \mu^{(1)}|\phi^{(0)}\rangle. \tag{22b}$$

Since $|\phi^{(0)}\rangle \neq 0$, but $\hat{L}^{(0)} = 0$, the first gives $\mu^{(0)} = 0$, leaving

$$\hat{L}^{(1)}|\phi^{(0)}\rangle = \mu^{(1)}|\phi^{(0)}\rangle = \frac{\partial \hat{\rho}_f^{(1)}}{\partial \lambda}|\phi^{(0)}\rangle. \tag{23}$$

Thus, to lowest order in the purity, a measurement that suffices to saturate the quantum CRB is one which is done in the eigenbasis of $\frac{\partial \hat{\rho}_f^{(1)}}{\partial \lambda}$.

IV. SINGLE-QUBIT, SINGLE-CHANNEL PROTOCOLS

A baseline against which to compare any metrology protocol is the SQSC protocol illustrated in Fig. 2(a). The analysis depends on whether the channel is unital or not and the results will be described in terms of the Bloch-sphere mapping of Eq. (2).

A. SQSC protocols for unital channels

For a single-qubit unital channel $\hat{\rho}_i^{(1)} := \hat{\sigma}_{\mathbf{r}_0}/2$, giving $\hat{\rho}_f^{(1)} = (M\mathbf{r}_0) \cdot \hat{\boldsymbol{\sigma}}/2$. This and Eqs. (20) and (21) imply that, *to lowest order in the purity,*

$$H = r^2 \, \mathbf{r}_0^\top \dot{M}^\top \dot{M} \mathbf{r}_0, \tag{24}$$

where the dot indicates the derivative with respect to the parameter. This forms a general result for SQSC protocols for unital channels.

Further analysis, all to lowest order only in the purity, uses the singular value decomposition for real matrices. Here $\dot{M} = ASB$, where A and B are each orthogonal 3×3 matrices and $S = s_1 P_1 + s_2 P_2 + s_3 P_3$ is a diagonal matrix with positive entries arranged so that $s_1 \geqslant s_2 \geqslant s_3$; here $\{P_i\}$ are projectors onto each of the three orthogonal directions associated with unit vectors $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3\}$. The orthogonality of A and projective nature of P_i implies that

$$H = r^2 \sum_{i=1}^{3} s_i^2 \mathbf{r}_0^\top B^\top P_i B \mathbf{r}_0. \tag{25}$$

Now $\mathbf{r}_0^\top B^\top P_i B \mathbf{r}_0 \geqslant 0$ and $\sum_{i=1}^{3} \mathbf{r}_0^\top B^\top P_i B \mathbf{r}_0 = 1$ implies that the *optimal lowest order SQSC protocol QFI* is

$$H_{\text{s opt}} = r^2 \, s_1^2. \tag{26}$$

This is attained with $\mathbf{r}_0 = B^\top \hat{\mathbf{e}}_1$, where $\hat{\mathbf{e}}_1$ is the unit vector associated with the maximum singular value in S. Note that, depending on the singular value decomposition, this might depend on the parameter to be estimated.

One measurement which can saturate the quantum CRB bound is a projective measurement onto the eigenbasis of

$\frac{\partial \hat{\rho}_f^{(1)}}{\partial \lambda}$. Here, for the optimal choice of input state, $\frac{\partial \hat{\rho}_f^{(1)}}{\partial \lambda} = (\dot{M}B^\top \hat{\mathbf{e}}_1) \cdot \hat{\boldsymbol{\sigma}}/2 = (AS\hat{\mathbf{e}}_1) \cdot \hat{\boldsymbol{\sigma}}/2$ and the resulting projective measurement operators are

$$\hat{\Pi}_\pm := \frac{1}{2} [\hat{I} \pm (A\hat{\mathbf{e}}_1) \cdot \hat{\boldsymbol{\sigma}}]. \tag{27}$$

Whenever the direction of $A\hat{\mathbf{e}}_1$ depends on the parameter, these projectors will also depend on the parameter to be estimated and adaptive measurement schemes [29] must be invoked to attain the QFI. But if the direction of $A\hat{\mathbf{e}}_1$ is independent of the parameter, then the method described here will yield a parameter-independent saturating measurement.

To summarize, with a unital channel subject to the SQSC protocol, the optimal QFI to lowest order in the purity is determined by finding the Bloch-sphere matrix M that represents the channel action and determining the singular value decomposition $\dot{M} = ASB$. The optimal QFI depends only on the maximal singular value s_1 and the protocol which attains this is to prepare the input state along the Bloch-sphere direction $B^\top \hat{\mathbf{e}}_1$, where $\hat{\mathbf{e}}_1$ is the direction associated with the maximal singular value in S, and then subject the qubit to the channel. One measurement that saturates the QCRB is a projection onto the Bloch-sphere direction $A\hat{\mathbf{e}}_1$.

Example: Unitary phase shift. The unitary phase shift about the z axis through angle λ is represented by $\hat{\rho}_i \mapsto \hat{\rho}_f = U^\dagger \hat{\rho}_i U$, where $U := e^{-i\lambda\hat{\sigma}_z/2}$. In the basis $\{\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}\}$,

$$M = \begin{pmatrix} \cos\lambda & -\sin\lambda & 0 \\ \sin\lambda & \cos\lambda & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{28}$$

Then

$$\dot{M} = \begin{pmatrix} -\sin\lambda & -\cos\lambda & 0 \\ \cos\lambda & -\sin\lambda & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{29}$$

which gives $S = \text{diag}(1, 1, 0)$ with various possibilities for A and B. The vector associated with the maximal singular value is any unit vector in the the xy plane. This gives an optimal lowest order QFI of $H_{\text{s opt}} = r^2$. The optimal QFI is attained using a state with Bloch-sphere input direction in the xy plane, for example $B^\top \hat{\mathbf{x}}$. The saturating measurement of Eq. (27) is a projective measurement along the direction $A\hat{\mathbf{x}}$. It is not possible that both the choice of initial Bloch-sphere direction and measurement direction can both be independent of the parameter; this is consistent with exact calculations.

Example: Phase-flip channel. The phase-flip channel maps $\hat{\rho}_i \mapsto \hat{\rho}_f = (1-\lambda)\hat{\rho}_i + \lambda\hat{\sigma}_z\hat{\rho}_i\hat{\sigma}_z$. In the basis $\{\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}\}$, $M = \text{diag}(1-2\lambda, 1-2\lambda, 1)$ and $\dot{M} = \text{diag}(-2, -2, 0)$ so that $S = \text{diag}(2, 2, 0)$ with $A = \text{diag}(-1, -1, 0)$ and $B = I$ as one possibility. This gives an optimal lowest order QFI is $H_{\text{s opt}} = 4r^2$, attained when the initial-state Bloch-sphere direction is in the xy plane. This agrees with approximations from the exact QFI [20]. The saturating measurement of Eq. (27) is a projection along the direction $\mathbf{r}_0$ and is parameter independent.

Example: Depolarizing channel. The depolarizing channel maps $\hat{\rho}_i \mapsto \hat{\rho}_f = (1-\lambda)\text{Tr}[\hat{\rho}_i]\hat{I} + \lambda\hat{\rho}_i$ and $M = \lambda I$ with $\dot{M} = I$. This indicates that the optimal lowest order QFI is $H_{\text{s opt}} = r^2$ and this is attained regardless of the choice of

initial-state vector. This is consistent with approximations for exact the QFI [21]. Again a saturating measurement from the SLD is parameter independent.

B. SQSC protocols for nonunital channels

For the more general nonunital channel acting on a single-qubit $\hat{I} \xrightarrow{\Gamma} \hat{I} + \hat{\sigma}_{\mathbf{d}}$ and thus

$$\hat{\rho}_{\mathrm{f}}^{(0)} = \frac{1}{2}(\hat{I} + \hat{\sigma}_{\mathbf{d}}). \qquad (30)$$

The resulting analysis, again all to lowest order in the purity, depends on whether $\mathbf{d}$ is parameter dependent. In Appendix C we show that if $\mathbf{d}$ is parameter dependent, then the zeroth order term in the QFI is generally nonzero and is

$$H_{\mathrm{s\,opt}} = \begin{cases} \dot{\mathbf{d}} \cdot \dot{\mathbf{d}} + \dfrac{1}{4(1-d^2)}\left[\dfrac{\partial d^2}{\partial \lambda}\right]^2 & \text{if } d \neq 1, \\ \dot{\mathbf{d}} \cdot \dot{\mathbf{d}} & \text{if } d = 1. \end{cases} \qquad (31)$$

Here $d := |\mathbf{d}|$. If $\mathbf{d}$ is parameter dependent then this is the optimal lowest order QFI for nonunital SQSC protocols. A key feature of such channels is that to lowest order in the purity, the QFI is independent of r and this could be attained by an input state with zero purity. A sufficient measurement that would attain this is a projective measurement onto the eigenbasis of the lowest order score operator $\hat{L}^{(0)}$ and is thus a measurement along the Bloch-sphere direction determined by $\dot{\mathbf{d}}$ (if $d = 1$) or $\dot{\mathbf{d}} + \frac{\partial \ln(1-d^2)}{\partial \lambda}\mathbf{d}$ (if $d \neq 1$).

On the other hand, if $\mathbf{d}$ is parameter independent then this will yield zero. Again as shown in Appendix C, if $\mathbf{d}$ is independent of the parameter then the lowest order term in the QFI is

$$H = r^2 \mathbf{r}_0^\top \left[\dot{M}^\top \dot{M} + \dfrac{d^2}{1-d^2}\dot{M}^\top P_{\hat{\mathbf{d}}}\dot{M}\right]\mathbf{r}_0, \qquad (32)$$

where $P_{\hat{\mathbf{d}}}$ is the projector onto the direction $\hat{\mathbf{d}}$. Note that if $d = 1$ then $M = 0$ and there is no parameter dependence to the channel at all. We can ignore this case. The entire operator within braces is positive and a singular value decomposition of this will eventually yield the optimal lowest order QFI and initial Bloch-sphere direction. Note that, comparing with Eq. (24), this indicates that channels with nonzero constant Bloch-sphere shift vector will typically enhance the estimation accuracy of the channel corresponding to the Bloch-sphere matrix M alone by effectively increasing the purity of the state.

Example: Generalized amplitude damping The generalized amplitude damping channel maps $\hat{\rho}_{\mathrm{i}} \mapsto \hat{\rho}_{\mathrm{f}} = \sum_{i=1}^{4} E_i^\dagger \hat{\rho}_{\mathrm{i}} E_i$, where

$$E_1 = \sqrt{p}\begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\lambda} \end{pmatrix},$$

$$E_2 = \sqrt{p}\begin{pmatrix} 0 & \sqrt{\lambda} \\ 0 & 0 \end{pmatrix},$$

$$E_3 = \sqrt{1-p}\begin{pmatrix} \sqrt{1-\lambda} & 0 \\ 0 & 1 \end{pmatrix},$$

$$E_4 = \sqrt{1-p}\begin{pmatrix} 0 & 0 \\ \sqrt{\lambda} & 0 \end{pmatrix}, \qquad (33)$$

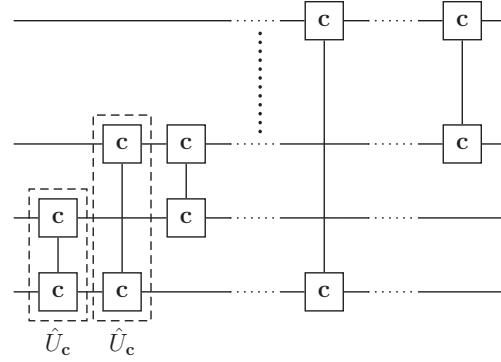with $0 \leqslant p \leqslant 1$.

FIG. 3. The preparatory unitary for a general symmetric pairwise correlated scheme considered in this article. The symbols within the blue dashed frame represent a single iteration of $\hat{U}_{\mathbf{c}}$; the boxes indicate the two qubits on which the gate acts.

Then [23] $M = \mathrm{diag}(\sqrt{1-\lambda}, \sqrt{1-\lambda}, 1-\lambda)$ and $\mathbf{d} = \lambda(2p-1)\hat{\mathbf{z}}$. This yields $H_{\mathrm{s\,opt}} = 1/[1 - \lambda^2(2p-1)^2]$. The optimal measurement that saturates the quantum CRB bound is a projective measurement along $\hat{\mathbf{z}}$.

Compiling these results gives in a complete characterization of the lowest order QFI terms for SQSC protocols for all channels: Eq. (26) for unital channels, Eq. (31) for nonunital channels with a parameter-dependent Bloch-sphere shift vector, and Eq. (32) for nonunital channels with a parameter-independent shift.

V. SYMMETRIC PAIRWISE CORRELATED PROTOCOLS

The central question is whether there is an entanglement-assisted protocol which can yield a larger QFI per channel invocation than the optimal SQSC protocol with the same purity. Previous results for parameter estimation for the phase-shift, phase-flip, and depolarizing channel showed that this is possible for a particular correlating preparatory unitary [18,20,21]. We consider a generalization of this for any qubit channel.

Specifically, we consider a protocol where the preparatory unitary is constructed from the two-qubit unitary

$$\hat{U}_{\mathbf{c}} := \frac{1}{2}(\hat{I} \otimes \hat{I} + \hat{I} \otimes \hat{\sigma}_{\mathbf{c}} + \hat{\sigma}_{\mathbf{c}} \otimes \hat{I} - \hat{\sigma}_{\mathbf{c}} \otimes \hat{\sigma}_{\mathbf{c}}), \qquad (34)$$

where $\mathbf{c}$ is a unit vector, which determines the *Bloch-sphere control direction* of this gate. The preparatory unitary is defined to consist of a product of such unitaries, one for each distinct pair of qubits, as illustrated in Fig. 3. This is symmetrical under interchange of qubits and only involves pairwise correlating unitaries; we term it a symmetric pairwise correlated protocol. If $\mathbf{c} = \hat{\mathbf{z}}$ this is the controlled-Z gate used along with other single-qubit Hadamard gates in the correlated-state protocols studied previously [20,21].

Aside from demonstrating gains in the past, protocols of this type are interesting because the number of basic two-qubit gates scales quadratically in the total number of qubits and in many physical settings these gates are relatively easily constructed. For example, in solution-state NMR implementations of quantum information processing they have been

implemented experimentally since the outset of that field [38–40].

Under this preparatory unitary, the two lowest order terms in the input state to the channels are $\hat{\rho}_i^{(0)} = \hat{I}^{\otimes n}/N$ and $\hat{\rho}_i^{(1)} = \hat{U}_{\text{prep}}\hat{\rho}_0^{(1)}\hat{U}_{\text{prep}}^\dagger$; these will be sufficient for determining the lowest order terms in the QFI. In Appendix D we show that

$$\hat{\rho}_i^{(1)} = \frac{1}{N}\left(\hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{c}}^{\otimes(n-1)} + \cdots + \hat{\sigma}_{\mathbf{c}}^{\otimes(n-1)} \otimes \hat{\sigma}_{\mathbf{r}_0}\right)$$
$$+ \frac{\mathbf{r}_0 \cdot \mathbf{c}}{N}\left(\hat{\sigma}_{\mathbf{c}} \otimes \hat{I}^{\otimes(n-1)} + \cdots + \hat{I}^{\otimes(n-1)} \otimes \hat{\sigma}_{\mathbf{c}}\right)$$
$$- \frac{n\,\mathbf{r}_0 \cdot \mathbf{c}}{N}\,\hat{\sigma}_{\mathbf{c}}^{\otimes n}. \tag{35}$$

Note that within the first parentheses, there are n different terms, each containing a single factor of $\hat{\sigma}_{\mathbf{r}_0}$. Similarly within the second parentheses there are also n different terms, each containing a single factor of $\hat{\sigma}_{\mathbf{c}}$.

We assume that the channel is invoked once on a single qubit. Again the analysis depends on whether the channel is unital or not.

A. Symmetric pairwise correlated protocol for unital channels with a single invocation

Assume that a unital channel acts once on the leftmost qubit in the tensor product representation. Then the terms in the first order term in the input state of Eq. (35) are mapped by the channel as

$$\hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{c}}^{\otimes(n-1)} \mapsto (M\mathbf{r}_0) \cdot \hat{\boldsymbol{\sigma}} \otimes \hat{\sigma}_{\mathbf{c}}^{\otimes(n-1)},$$
$$\hat{\sigma}_{\mathbf{c}} \otimes \hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{c}}^{\otimes(n-2)} \mapsto (M\mathbf{c}) \cdot \hat{\boldsymbol{\sigma}} \otimes \hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{c}}^{\otimes(n-2)},$$
$$\hat{\sigma}_{\mathbf{c}} \otimes \hat{I}^{\otimes(n-1)} \mapsto (M\mathbf{c}) \cdot \hat{\boldsymbol{\sigma}} \otimes \hat{I}^{\otimes(n-1)},$$
$$\hat{I} \otimes \hat{\sigma}_{\mathbf{c}} \otimes \hat{I}^{\otimes(n-2)} \mapsto \hat{I} \otimes \hat{\sigma}_{\mathbf{c}} \otimes \hat{I}^{\otimes(n-2)},$$
$$\hat{\sigma}_{\mathbf{c}}^{\otimes(n)} \mapsto (M\mathbf{c}) \cdot \hat{\boldsymbol{\sigma}} \otimes \hat{\sigma}_{\mathbf{c}}^{\otimes(n-1)}, \tag{36}$$

where M is the channel Bloch-sphere matrix. Thus

$$\hat{\rho}_f^{(1)} = \frac{1}{N}\Big[(\dot{M}\mathbf{r}_0) \cdot \hat{\boldsymbol{\sigma}} \otimes \hat{\sigma}_{\mathbf{c}}^{\otimes(n-1)}$$
$$+ (\dot{M}\mathbf{c}) \cdot \hat{\boldsymbol{\sigma}} \otimes \hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{c}}^{\otimes(n-2)}$$
$$+ \cdots + (\dot{M}\mathbf{c}) \cdot \hat{\boldsymbol{\sigma}} \otimes \hat{\sigma}_{\mathbf{c}}^{\otimes(n-2)} \otimes \hat{\sigma}_{\mathbf{r}_0}\Big]$$
$$+ \frac{\mathbf{r}_0 \cdot \mathbf{c}}{N}(\dot{M}\mathbf{c}) \cdot \hat{\boldsymbol{\sigma}} \otimes \hat{I}^{\otimes(n-1)}$$
$$- \frac{n\mathbf{r}_0 \cdot \mathbf{c}}{N}(\dot{M}\mathbf{c}) \cdot \hat{\boldsymbol{\sigma}} \otimes \hat{\sigma}_{\mathbf{c}}^{\otimes(n-1)}. \tag{37}$$

This yields our main result (see Appendix E for a proof) for unital channels: if the channel is invoked once on a single qubit when the symmetric pairwise correlated protocol is used then to the lowest order in the purity (remaining analysis is all to lowest order)

$$H = r^2\mathbf{r}_0^\top[(I - P_{\mathbf{c}})\dot{M}^\top\dot{M}(I - P_{\mathbf{c}}) + (2 - n)P_{\mathbf{c}}\dot{M}^\top\dot{M}P_{\mathbf{c}}]\mathbf{r}_0$$
$$+ r^2(n-1)\mathbf{c}^\top\dot{M}^\top\dot{M}\mathbf{c}, \tag{38}$$

where I is the 3×3 identity matrix and $P_{\mathbf{c}}$ is the projector onto the control direction vector $\mathbf{c}$.

For a given channel there remains the task of choosing the control direction vector and initial-state Bloch-sphere vector

so as to maximize the QFI of Eq. (38). The details depend on the channel, but whenever a unital channel is invoked once it is possible to bound the lowest nonzero term in the QFI. As shown in Appendix F,

$$nr^2s_1^2 - r^2s_1^2\left(1 - \frac{s_2^2}{s_1^2}\right) \leqslant H \leqslant nr^2s_1^2, \tag{39}$$

where $s_1 \geqslant s_2 \geqslant s_3 \geqslant 0$ are the singular values of $\dot{M}$. The upper bound can only be saturated when $\mathbf{r}_0$ and $\mathbf{c}$ are perpendicular. The lower bound is attained for a particular choice of perpendicular $\mathbf{r}_0$ and $\mathbf{c}$ (see Appendix F for details).

Equations (26) and (39) allow for comparison of the symmetric pairwise correlated protocol against the SQSC protocol to lowest order for unital channels. Here

$$\left[n - \left(1 - \frac{s_2^2}{s_1^2}\right)\right]H_{\text{s opt}} \leqslant H_{\text{corr opt}} \leqslant nH_{\text{s opt}}, \tag{40}$$

where $H_{\text{corr opt}}$ is the optimal QFI for the correlated protocol over all choices of $\mathbf{c}$ and $\mathbf{r}_0$. Thus

$$n - \left(1 - \frac{s_2^2}{s_1^2}\right) \leqslant \frac{H_{\text{corr opt}}}{H_{\text{s opt}}} \leqslant n. \tag{41}$$

Since $s_2 \leqslant s_1$ this means that for large n and to lowest order in the purity, the symmetric pairwise correlated protocol roughly gives an n-fold gain over the SQSC protocol for any unital channel.

Sometimes a precise statement about the optimal QFI for this correlated-state protocol can be made. If $s_1 = s_2$, as is true for several commonly considered channels, the the two bounds of Eq. (39) are identical and $H_{\text{corr opt}} = nH_{\text{s opt}}$. As another example, if $s_2 = s_3 = 0$, the analysis of Appendix G shows that $H_{\text{corr opt}} = (n - 1)H_{\text{s opt}}$ and this is attained when $\mathbf{c}$ and $\mathbf{r}_0$ are perpendicular.

The realm of applicability of the bounds of Eqs. (39) and (41) can be assessed via higher order terms in the QFI. In Appendix H we show that if $\mathbf{r}_0$ and $\mathbf{c}$ are perpendicular, then $H^{(3)} = 0$ and that generally $H^{(4)}$ is of order n^2. Thus the fourth order contribution to the QFI scales as n^2r^4. Given that the third order contribution is zero and that the second order contribution scales as nr^2, it is clear that approximating the QFI via the lowest order nonzero contribution is valid only when $nr^2 \ll 1$.

B. Measurements for symmetric pairwise correlated protocol for unital channels

The remaining issue with this optimal protocol is to find a QRB saturating measurement. A projective measure in the eigenbasis of $\hat{\rho}_f^{(1)}$ suffices. For the optimal symmetric pairwise correlated protocol,

$$\hat{\rho}_f^{(1)} = \frac{1}{N}\Big[(\dot{M}\mathbf{r}_0) \cdot \hat{\boldsymbol{\sigma}} \otimes \hat{\sigma}_{\mathbf{c}}^{\otimes(n-1)}$$
$$+ (\dot{M}\mathbf{c}) \cdot \hat{\boldsymbol{\sigma}} \otimes \hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{c}}^{\otimes(n-2)} + \cdots$$
$$+ (\dot{M}\mathbf{c}) \cdot \hat{\boldsymbol{\sigma}} \otimes \hat{\sigma}_{\mathbf{c}}^{\otimes(n-2)} \otimes \hat{\sigma}_{\mathbf{r}_0}\Big]. \tag{42}$$

Sometimes this eigenbasis will depend on the parameter value, thus suggesting a measurement that would require knowledge of the parameter; this is a general issue which has
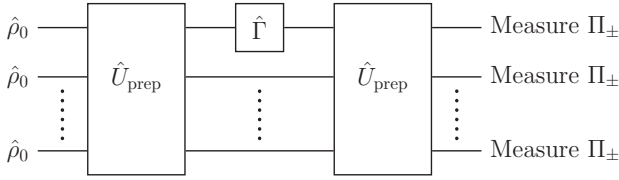
FIG. 4. Symmetric pairwise correlated scheme followed by local projective measurements. The measurements are chosen so that the associated projection operators are $\hat{\Pi}_{\pm} = (\hat{I} \pm \hat{\sigma}_{\mathbf{r}_0})/2$, where $\mathbf{r}_0$ is the initial state Bloch-sphere direction.

been addressed elsewhere [28,29]. Nonetheless the resulting measurement will always saturate the QFI.

A separate issue is whether there exists a series of local, single-qubit measurements that can yield the QFI. Whether such a procedure exists immediately after channel invocation is not clear. However, in Appendix I we show that if the channel invocation is following by another invocation of the preparatory unitary and this is followed by a local measurement on each qubit with appropriate choices of Bloch-sphere directions, then the resulting classical Fisher information equals the lower bound of Eq. (39). This merely entails another quadratic cost in terms of two-qubit gates. Figure 4 illustrates this scheme.

Specifically if, after the second invocation of the preparatory unitary, a local projective measurement corresponding to operators $\hat{\Pi}_{\pm} = (\hat{I} \pm \hat{\sigma}_{\mathbf{r}_0})/2$ is enacted on each qubit, where the Bloch-sphere control and initial state vectors satisfy $\mathbf{c} = B^{\top}\hat{\mathbf{e}}_1$ and $\mathbf{r}_0 = B^{\top}\hat{\mathbf{e}}_2$ where $\hat{\mathbf{e}}_1$ and $\hat{\mathbf{e}}_2$ are orthogonal unit vectors associated with the singular value decomposition of $\dot{M}$, then the classical Fisher information is

$$F = nr^2 s_1^2 - r^2 s_1^2 \left(1 - \frac{s_2^2}{s_1^2}\right). \tag{43}$$

This is exactly the lower bound of Eq. (39).

Thus it is always possible to attain the gain with a factor of at least $n - 1$ in estimation accuracy, using a local measurement scheme preceded by the preparatory unitary. We assess various important examples.

Example: Phase shift. For the phase-shift channel about the z axis, $s_1 = s_2 = 1$, $s_3 = 0$, and one possible choice of principal axis directions is $\hat{\mathbf{e}}_1 = \hat{\mathbf{x}}$, $\hat{\mathbf{e}}_2 = \hat{\mathbf{y}}$, and $\hat{\mathbf{e}}_3 = \hat{\mathbf{z}}$. By Eq. (39) this gives an optimal QFI, $H_{\text{corr opt}} = nH_{\text{s opt}}$. Taking $B = I$ in the singular value decomposition, the choices $\mathbf{c} = \hat{\mathbf{e}}_1$ and $\mathbf{r}_0 = \hat{\mathbf{e}}_2$ attain this. Since $s_1 = s_2$, the measurement scheme of Fig. 4 yields a classical Fisher information $F = nH_{\text{s opt}}$ that saturates the optimal QFI with a parameter-independent measurement.

Example: Phase flip. Here $s_1 = s_2 = -2$, $s_3 = 0$ with $\hat{\mathbf{e}}_3 = \hat{\mathbf{z}}$ and $\hat{\mathbf{e}}_1$ and $\hat{\mathbf{e}}_2$ any perpendicular unit vectors in the xy plane. This gives an optimal QFI, $H_{\text{corr opt}} = nH_{\text{s opt}}$. This is attained when $\mathbf{c} = \hat{\mathbf{e}}_1$ and $\mathbf{r}_0 = \hat{\mathbf{e}}_2$. Again $s_1 = s_2$ and the measurement scheme of Fig. 4 yields a classical Fisher information $F = nH_{\text{s opt}}$ with a parameter-independent choice of measurement. This agrees with lowest order approximations from exact expressions for all purities [20].

Example: Depolarizing channel. Here $s_1 = s_2 = s_3 = 1$, giving an optimal QFI, an optimal QFI, $H_{\text{corr opt}} = nH_{\text{s opt}}$,

which is attained when $\mathbf{c} = \hat{\mathbf{e}}_1$ and $\mathbf{r}_0 = \hat{\mathbf{e}}_2$ are any two orthogonal vectors. Here also $s_1 = s_2$ and the measurement scheme of Fig. 4 yields a classical Fisher information $F = nH_{\text{s opt}}$ with a parameter-independent choice of measurement. This agrees with lowest order approximations from exact expressions for all purities [21].

C. Gains for symmetric pairwise correlated protocol for unital channels

These examples illustrate the general approximately n-fold gain offered by the symmetric pairwise correlated protocol whenever $n \ll 1/r^2$. This is reminiscent of the gains described by the Heisenberg limit, where the QFI scales as n^2, over the standard quantum limit, where the QFI scales as n, and n is the number of probes or channel invocations [5]. One difference is that in the protocol of this article the channel only acts once on one of the probes, whereas the typical n-fold Heisenberg scaling gain involves multiple copies of the channel or multiple systems subjected to the same channel. Mathematically the gain in our protocol arises from the fact that when the control and initial state Bloch-sphere vectors are perpendicular the channel input state of Eq. (35) has n terms on which channel can have a nontrivial effect while, in the absence of the preparatory unitary, the initial state of Eq. (12) only offers one term on which the channel has a nontrivial effect. Somehow, the preparatory unitary of the symmetric pairwise correlated protocol has produced correlations that distribute information amongst the qubits so as to effectively mimic action of the channel as though it has acted on every qubit. It has also managed to do this in a way which works for all unital channels although the control and initial state Bloch-sphere directions will need to be adjusted depending on the channel.

Gains of this type in quantum metrology are often associated with the use of entangled states and it may be asked whether this is responsible for the gains here. However, as shown previously for the phase-shift [18], phase-flip [20], and depolarizing channels [21], for the two-qubit case the system state are separable whenever the purity satisfies $r < \sqrt{2} - 1$. This rules out entanglement as a source of the gains presented in this article. On the other hand, the same studies did show that for the two-qubit state, the quantum discord is nonzero for all purities, $r > 0$, which encompass those of this article. How these statements might apply beyond two qubits is currently unclear.

D. Symmetric pairwise correlated protocol for nonunital channels

For nonunital channels where $\mathbf{d}$ depends on the parameter, the lowest nonzero term in the QFI is the zeroth order term. The zeroth order term in the density operator in the symmetric pairwise correlated-state protocol will be the same as that in the SQSC protocol; this is evident from setting $r = 0$ in the formalism. Thus for such channels, the lowest order term in QFI in the symmetric pairwise correlated protocol is the same as that for the SQSC protocol. To lowest order there is nothing to gain from the correlated-state protocol for such cases.

A possible explanation for this is that, for such nonunital channels, even the maximally mixed state would be fruitful for

$$\boxed{\begin{array}{c} \text{Single qubit channel:} \\[4pt] \hat{\rho}_{\text{i}} = \frac{1}{2}\left(\hat{I} + r\hat{\sigma}_{\hat{r}_{\text{i}}}\right) \mapsto \frac{1}{2}\left(\hat{I} + \hat{\sigma}_{\mathbf{r}_{\text{f}}}\right) \end{array}}$$

$$\boxed{\begin{array}{c} \text{Channel described by Bloch-sphere map:} \\[4pt] \mathbf{r}_{\text{f}} = rM\hat{r}_{\text{i}} + \mathbf{d} \end{array}} \longrightarrow \boxed{\begin{array}{c} \text{Singular values of } \dot{M}: \\[4pt] s_1 \geqslant s_2 \geqslant s_3 \geqslant 0 \end{array}}$$

$$\boxed{\begin{array}{c} \textbf{Non-unital channel with } \dot{\mathbf{d}} \neq 0\text{:} \\[4pt] \text{Optimal SQSC protocol QFI:} \\[4pt] H_{\text{s opt}} = \begin{cases} \dot{\mathbf{d}} \cdot \dot{\mathbf{d}} + \dfrac{1}{4(1-d^2)}\left[\dfrac{\partial d^2}{\partial \lambda}\right]^2 & \text{if } d \neq 1 \\[10pt] \dot{\mathbf{d}} \cdot \dot{\mathbf{d}} & \text{if } d = 1. \end{cases} \\[6pt] \text{Optimal correlated protocol QFI satisfies:} \\[4pt] H_{\text{corr opt}} = H_{\text{s opt}} \\[4pt] \textbf{Correlated protocol offers no gain.} \end{array}}$$
$$\boxed{\begin{array}{c} \textbf{Unital channel } (\mathbf{d} = 0)\text{:} \\[4pt] \text{Optimal SQSC protocol QFI:} \\[4pt] H_{\text{s opt}} = r^2 s_1^2 \\[6pt] \text{Optimal correlated protocol QFI satisfies:} \\[4pt] n - \left(1 - \dfrac{s_2^2}{s_1^2}\right) \leqslant \dfrac{H_{\text{corr opt}}}{H_{\text{s opt}}} \leqslant n \\[6pt] \textbf{Correlated protocol offers gain by a} \\ \textbf{factor of at least } n-1. \end{array}}$$
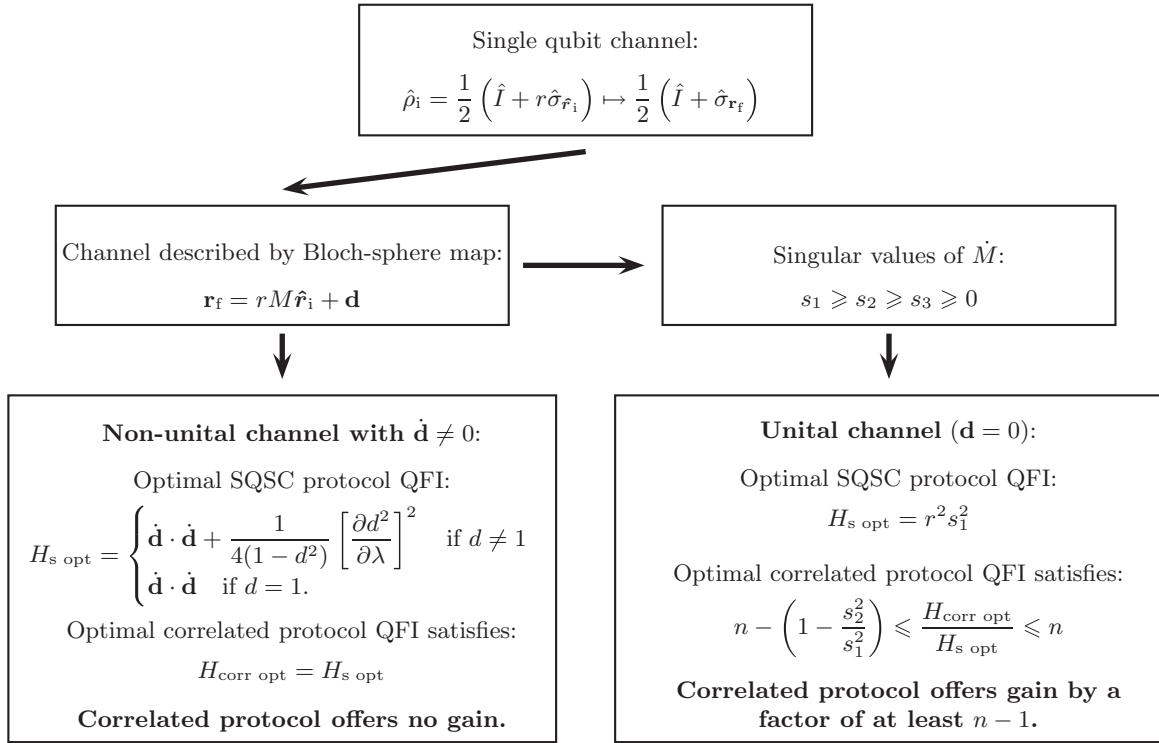
FIG. 5. Summary of the key results for single-qubit channel parameter estimation. Results are all to lowest nonzero order in purity. The correlated protocol refers to a symmetric pairwise correlated-state protocol with the channel invoked once on one of n qubits.

estimation as the channel maps $\hat{I} \overset{\Gamma}{\mapsto} \hat{I} + \hat{\sigma}_{\mathbf{d}}$ and the parameter could be estimated from the final system state. A maximally mixed initial state would be unaffected by any preparatory unitary and thus any correlated-state protocol would not make a difference for estimation if this were the system's initial state. The lowest order analysis that we have used only retains this maximally mixed term and therefore we might not have expected any gains from the correlated-state protocol here.

VI. SUMMARY OF RESULTS

We summarize the main results for the single-qubit single-channel (SQSC) protocol and the symmetric pairwise correlated protocols (see Figs. 2 and 3 and the associated descriptions). These assume that all qubits are initially in the state $\hat{\rho}_0 = (\hat{I} + r\hat{\sigma}_{\mathbf{r}_0})/2$ and the results apply whenever the purity satisfies $nr^2 \ll 1$. Figure 5 provides a schematic that describes the main results in terms of the relevant quantum Fisher information for single-qubit channels.

VII. CONCLUSION

We have compared quantum parameter estimation protocols for qubit channels when the available states are mixed with very low purity and where the channel is invoked once. We have shown that for any unital channel, with initial-state purity r, the particular n qubit correlated input state generated by the symmetric pairwise correlated protocol provides a roughly n-fold increase in estimation accuracy over protocols that use uncorrelated states provided that $nr^2 \ll 1$. These results agree with approximations from exact results

for the known special cases of the phase-shift [18], phase-flip [20], and depolarizing channels [21]. We also presented a measurement scheme that requires one more application of the preparatory unitary followed by local single-qubit measurements that yields a classical Fisher information which saturates the lower bound in the QFI and this gives the roughly n-fold gain in accuracy. This is parameter independent. There still remains the issue of finding generic measurement choices that are independent of the parameter and that yield a classical Fisher information which exactly saturates the optimal QFI.

For nonunital channels with a parameter-dependent shift, to lowest order in the purity, there is no improvement in estimation accuracy using these particular parameters; this provides a first glimpse into amplitude-damping channel parameter with nonpure initial states.

The formalism used here could be extended to situations where the channel is invoked on more than one of the qubits. The first order term in the channel input density operator would still have the form of Eq. (35) but, for unital channels, the channel actions would produce an expression analogous to those of Eq. (37) with more than one factor of the Bloch-sphere matrix appearing in each term. Differentiation would then yield expressions for the QFI involving both the Bloch-sphere matrix and its derivative. This would require modifying the singular value analysis used in this article. Additionally, it is already known for certain unital channels [20,21] that these protocols only offer advantages over a subset of possible parameter values and thus the universal results will not apply. Studies into restricted regions of the parameter space, where these protocols could offer advantages would still be warranted.

The symmetric pairwise correlated-state protocol considered here generalizes those used previously and for unital channels. This is some improvement over previous studies [18,20,21] in estimation with noisy states, but even here the scheme awaits optimization over choices of initial-state Bloch-sphere and control vector directions. It is also possible that another type of preparation scheme might be optimal. The structure of the lowest order terms in the density operator in this formalism give some insight into the origin of the increase in the QFI. After the preparatory unitary, every qubit provides a term in the expression for the system state such that the channel has a nontrivial action on this term. This might be able to yield some insights into the origins of the accuracy enhancement and possible ways to improve it.

Finally, we note that there have been many studies of parameter estimation, typically of unitary parameters, in the presence of noisy processes (for example [9,32,41,42]). Such studies consider the situation where noise appears during or after action of the channel whose parameter is to be estimated and usually yield bounds on the QFI that depend on noise parameters. For example, in [32] a general bound was developed and applied to optical phase estimation where noise was introduced during the phase evolution. The framework presented there resulted in a QFI which, when sufficient noise is present, scales as the number of probes n and does not yield an n-fold advantage over independent channel invocation protocols. However, although this study presents a general technique, these scaling results appear for the specific situation where the state space for the system is infinite dimensional. Our initial attempts to use the general framework presented there for qubit systems where the noise is introduced at the outset have not yielded instructive bounds. In part, this must be due to the choice of representation of channel that also includes initial noise. However, the framework of [32] omits a specification of the optimal choice of such a channel. The exact relationship between this framework and the situation where noise is initially present in qubit channels warrants further investigation.

APPENDIX A: CHANNEL PARAMETRIZATION

Consider a channel whose effect is described by Eq. (2). First, $|\mathbf{r}_f| \leqslant 1$ for all possible inputs and for $r = 0$ this implies $|\mathbf{d}| \leqslant 1$. We argue that if $M \neq 0$, then $|\mathbf{d}| \neq 1$. To do so, $|\mathbf{r}_f|^2 = r^2 \mathbf{r}_i^\top M^\top M \mathbf{r}_i + |\mathbf{d}|^2 + r r_i^\top M^\top \mathbf{d} + r \mathbf{d}^\top M \mathbf{r}_i$, where $\top$ indicates the transpose. Assuming that $|\mathbf{d}| = 1$, taking the case where $r = 1$ and noting that the first term is positive this implies that for any unit vector $\mathbf{r}_i$, one of the terms $\mathbf{d}^\top M \mathbf{r}_i$ or $\mathbf{r}_i^\top M^\top \mathbf{d} = (\mathbf{d}^\top M \mathbf{r}_i)^\top$ must be negative. Thus for any unit vector $\mathbf{r}_i$, $\mathbf{d}^\top M \mathbf{r}_i$ must be neg-

ative. Now the singular value decomposition implies that $M = ASB$ where A and B are orthogonal matrices and S is diagonal with positive entries along the diagonal. Thus for any $\mathbf{r}_i$, $\mathbf{d}^\top ASB\mathbf{r}_i$ must be negative. Letting $\mathbf{d}' := A^\top \mathbf{d}$ and $\mathbf{r}_i' := B\mathbf{r}_i$ gives that $\mathbf{d}^\top ASB\mathbf{r}_i = \sum_j d_j' s_j r_j'$, where $s_j \geqslant 0$ are the singular values of M.

Thus, regardless $\mathbf{d}'$ there will always be some choices of $\mathbf{r}_i'$ such that this is positive. This implies that, regardless of of $\mathbf{d}$, there will be some choices of $\mathbf{r}_i$ so that $\mathbf{d}^\top M \mathbf{r}_i$ is positive. Thus, $|\mathbf{d}| = 1$ is only possible when $M = 0$.

APPENDIX B: ESCHER FRAMEWORK FOR PHASE-FLIP CHANNEL PARAMETER ESTIMATION WITH NOISY INITIAL STATES

We consider estimating the parameter λ in a phase-flip channel that acts on a single qubit according to $\hat{\rho}_i \mapsto (1 - \lambda)\hat{\rho}_i + \lambda \hat{\sigma}_z \hat{\rho}_i \hat{\sigma}_z$. The initial state for an individual qubits $\hat{\rho}_0 = (\hat{I} + r\hat{\sigma}_{\mathbf{r}_0})/2$ can be generated from a pure initial state $\rho_p = (\hat{I} + \hat{\sigma}_{\mathbf{r}_0})/2$ by a depolarizing channel with depolarizing parameter r. Denote the Kraus operators that generate the product of n such mixed states from n pure states by $\{\hat{\Pi}_{dj}(r)\}$. The preparatory unitary $\hat{U}_{prep}$ generates the channel input state $\hat{\rho}_i = \sum_j \hat{U}_{prep} \hat{\Pi}_{dj} \rho_p^{\otimes n} \hat{\Pi}_{dj}^\dagger \hat{U}_{prep}^\dagger$. Then denote the Kraus operators for the channel by $\{\hat{\Pi}_{chk}(\lambda)\}$. Thus the final premeasurement state is $\hat{\rho}_f = \sum_{j,k} \hat{\Pi}_{chk} \hat{U}_{prep} \hat{\Pi}_{dj} \rho_p^{\otimes n} \hat{\Pi}_{dj}^\dagger \hat{U}_{prep}^\dagger \hat{\Pi}_{chk}^\dagger$. Together these give *one set* of Kraus operators $\{\hat{\Pi}_{chk}(\lambda)\hat{U}_{prep}\hat{\Pi}_{dj}(r)\}$ that map the pure initial state to the final state produced by the channel. It follows that the operators used in the Escher framework are

$$
\begin{aligned}
\hat{H}_1 &= \sum_{j,k} \frac{\partial}{\partial \lambda} (\hat{\Pi}_{chk}(\lambda)\hat{U}_{prep}\hat{\Pi}_{dj}(r))^\dagger \\
&\quad \times \frac{\partial}{\partial \lambda} (\hat{\Pi}_{chk}(\lambda)\hat{U}_{prep}\hat{\Pi}_{dj}(r)) \\
&= \sum_{j,k} \hat{\Pi}_{dj}^\dagger \hat{U}_{prep}^\dagger \frac{\partial \hat{\Pi}_{chk}^\dagger}{\partial \lambda} \frac{\partial \hat{\Pi}_{chk}}{\partial \lambda} \hat{U}_{prep} \hat{\Pi}_{dj}.
\end{aligned} \tag{B1}
$$

Then straightforward algebra gives

$$
\text{Tr}(\hat{H}_1 \rho_p) = \text{Tr}\left(\sum_k \frac{\partial \hat{\Pi}_{chk}^\dagger}{\partial \lambda} \frac{\partial \hat{\Pi}_{chk}}{\partial \lambda} \hat{\rho}_i\right), \tag{B2}
$$

where $\hat{\rho}_i$ is the channel input state generated by the preparatory unitary from the noisy initial states. Similarly

$$
\text{Tr}(\hat{H}_2 \rho_p) = i \text{Tr}\left(\sum_k \frac{\partial \hat{\Pi}_{chk}^\dagger}{\partial \lambda} \hat{\Pi}_{chk} \hat{\rho}_i\right). \tag{B3}
$$

Now one possible Kraus representation for a single-qubit phase-flip channel is $\hat{\Pi}_{ch\,1} = \sqrt{1 - \lambda}\,\hat{I}$ and $\hat{\Pi}_{ch\,2} = \sqrt{\lambda}\,\hat{\sigma}_z$. This yields, for a single-qubit channel,

$$
\sum_k \frac{\partial \hat{\Pi}_{chk}^\dagger}{\partial \lambda} \frac{\partial \hat{\Pi}_{chk}}{\partial \lambda} = \frac{1}{4\lambda(1 - \lambda)}\,\hat{I}, \tag{B4a}
$$

$$
\sum_k \frac{\partial \hat{\Pi}_{chk}^\dagger}{\partial \lambda} \hat{\Pi}_{chk} = 0. \tag{B4b}
$$

Consider now the situation where only one qubit is present and the phase-flip channel is invoked once. Then Eqs. (8) and (B4) bound the QFI via $H \leqslant 1/\lambda(1 - \lambda)$ and this does not even refer to the initial-state purity. However, the known optimal single-qubit QFI [20] for the phase-flip channel acting on a single qubit with purity r is $H = 4r^2/[1 - (1 - 2\lambda)^2 r^2]$. Straightforward algebra shows that the bound given by the Escher framework and the particular choice of Kraus representation is strictly larger than the known optimal QFI whenever $r < 1$.

APPENDIX C: QFI FOR SQSC PROTOCOLS FOR NONUNITAL CHANNELS

Regardless of whether **d** depends on the parameter or not, the lowest order version of Eq. (15) gives $\hat{\rho}_{\mathrm{f}}^{(0)} = [\hat{L}^{(0)}\hat{\rho}_{\mathrm{f}}^{(0)} + \hat{\rho}_{\mathrm{f}}^{(0)}\hat{L}^{(0)}]/2$. The algebraic method offered in [20] gives

$$\hat{L}^{(0)} = \frac{1}{2} \frac{\partial \ln (1 - d^2)}{\partial \lambda} \hat{I} + \left[\dot{\mathbf{d}} - \frac{1}{2} \frac{\partial \ln (1 - d^2)}{\partial \lambda} \mathbf{d} \right] \cdot \hat{\boldsymbol{\sigma}}$$

$$\tag{C1}$$

when $d := |\mathbf{d}| \neq 1$ and

$$\hat{L}^{(0)} = \hat{\sigma}_{\mathbf{d}} \tag{C2}$$

if $d = 1$.

Direct substitution into the lowest order version of Eq. (16) yields the result of Eq. (31).

If **d** is parameter independent then $\hat{\rho}_{\mathrm{f}}^{(0)} = 0$ and $\hat{L}^{(0)} = 0$ and, as before, $H^{(0)} = H^{(1)} = 0$. The next order term is attained via Eq. (15), which with Eq. (30) gives

$$\hat{\rho}_{\mathrm{f}}^{(1)} = \frac{1}{2} \left[\hat{L}^{(1)}\hat{\rho}_{\mathrm{f}}^{(0)} + \hat{\rho}_{\mathrm{f}}^{(0)}\hat{L}^{(1)} \right]$$
$$= \frac{1}{2} \hat{L}^{(1)} + \frac{1}{4} \left[\hat{L}^{(1)}\hat{\sigma}_{\mathbf{d}} + \hat{\sigma}_{\mathbf{d}}\hat{L}^{(1)} \right]. \tag{C3}$$

If $d \neq 1$, then the solution to this, which can be verified by direct substitution, is

$$\hat{L}^{(1)} = \frac{2 - d^2}{1 - d^2} \hat{\rho}_{\mathrm{f}}^{(2)} - \frac{1}{1 - d^2} \left[\hat{\rho}_{\mathrm{f}}^{(1)}\hat{\sigma}_{\mathbf{d}} + \hat{\sigma}_{\mathbf{d}}\hat{\rho}_{\mathrm{f}}^{(1)} \right]$$

$$+ \frac{1}{1 - d^2} \left[\hat{\sigma}_{\mathbf{d}}\hat{\rho}_{\mathrm{f}}^{(1)}\hat{\sigma}_{\mathbf{d}} \right]. \tag{C4}$$

The remaining case where **d** is a parameter-independent unit vector requires $M = 0$ and this leaves no parameter dependence. We can ignore this. Equations (14) and (C4) give that if **d** is parameter independent, then to lowest order in the purity,

$$H = r^2 \frac{2 - d^2}{1 - d^2} \operatorname{Tr} \left[(\hat{\rho}_{\mathrm{f}}^{(1)})^2 \right] - r^2 \frac{2}{1 - d^2} \operatorname{Tr} \left[(\hat{\rho}_{\mathrm{f}}^{(1)})^2 \hat{\sigma}_{\mathbf{d}} \right]$$

$$+ r^2 \frac{1}{1 - d^2} \operatorname{Tr} \left[\hat{\sigma}_{\mathbf{d}}\hat{\rho}_{\mathrm{f}}^{(1)}\hat{\sigma}_{\mathbf{d}}\hat{\rho}_{\mathrm{f}}^{(1)} \right]. \tag{C5}$$

Then $\hat{\rho}_{\mathrm{f}}^{(1)} = (M\mathbf{r}_0) \cdot \hat{\boldsymbol{\sigma}}/2$ gives

$$H = r^2 \mathbf{r}_0^\top \dot{M}^\top \dot{M} \mathbf{r}_0 + \frac{r^2}{1 - d^2} (\mathbf{d}^\top \dot{M} \mathbf{r}_0)^2, \tag{C6}$$

and this yields the result of Eq. (32).

APPENDIX D: CHANNEL INPUT STATES

Computing the effects of the preparatory unitary on the entails algebra of Pauli operators. Explicit multiplication show that, for any vectors **a** and **b**,

$$\hat{\sigma}_{\mathbf{a}}\hat{\sigma}_{\mathbf{b}} = \mathbf{a} \cdot \mathbf{b}\,\hat{I} + i(\mathbf{a} \times \mathbf{b}) \cdot \hat{\boldsymbol{\sigma}}. \tag{D1}$$

The effects of the preparatory unitary can be determined by via the action of $\hat{U}_{\mathbf{c}}$ on pairwise products of operators. We will show that

$$\hat{U}_{\mathbf{c}}(\hat{\sigma}_{\mathbf{a}} \otimes \hat{I})\hat{U}_{\mathbf{c}}^\dagger = \hat{\sigma}_{\mathbf{a}} \otimes \hat{\sigma}_{\mathbf{c}}$$
$$+ (\mathbf{a} \cdot \mathbf{c})(\hat{\sigma}_{\mathbf{c}} \otimes \hat{I} - \hat{\sigma}_{\mathbf{c}} \otimes \hat{\sigma}_{\mathbf{c}}). \tag{D2}$$

To demonstrate (D2), consider first $(\hat{\sigma}_{\mathbf{a}} \otimes \hat{I})\hat{U}_{\mathbf{c}}^\dagger$. Noting that $\hat{U}_{\mathbf{c}}^\dagger = \hat{U}_{\mathbf{c}}$, repeatedly using (D1) and the fact that **c** is a unit vector gives

$$(\hat{\sigma}_{\mathbf{a}} \otimes \hat{I})\hat{U}_{\mathbf{c}}^\dagger = \frac{1}{2} [\mathbf{a} \cdot \mathbf{c}\,\hat{I} \otimes (\hat{I} - \hat{\sigma}_{\mathbf{c}}) + \hat{\sigma}_{\mathbf{a}} \otimes (\hat{I} + \hat{\sigma}_{\mathbf{c}})$$
$$+ i\hat{\sigma}_{\mathbf{a} \times \mathbf{c}} \otimes (\hat{I} - \hat{\sigma}_{\mathbf{c}})]. \tag{D3}$$

Then multiplying this separately by each term in $\hat{U}_{\mathbf{c}}$ gives

$$\frac{1}{2}\hat{I} \otimes \hat{I}(\hat{\sigma}_{\mathbf{a}} \otimes \hat{I})\hat{U}_{\mathbf{c}}^\dagger = \frac{1}{4}[\mathbf{a} \cdot \mathbf{c}\,\hat{I} \otimes (\hat{I} - \hat{\sigma}_{\mathbf{c}})$$
$$+ \hat{\sigma}_{\mathbf{a}} \otimes (\hat{I} + \hat{\sigma}_{\mathbf{c}})$$
$$+ i\hat{\sigma}_{\mathbf{a} \times \mathbf{c}} \otimes (\hat{I} - \hat{\sigma}_{\mathbf{c}})], \tag{D4a}$$

$$\frac{1}{2}\hat{I} \otimes \hat{\sigma}_{\mathbf{c}}(\hat{\sigma}_{\mathbf{a}} \otimes \hat{I})\hat{U}_{\mathbf{c}}^\dagger = \frac{1}{4}[-\mathbf{a} \cdot \mathbf{c}\,\hat{I} \otimes (\hat{I} - \hat{\sigma}_{\mathbf{c}})$$
$$+ \hat{\sigma}_{\mathbf{a}} \otimes (\hat{I} + \hat{\sigma}_{\mathbf{c}})$$
$$- i\hat{\sigma}_{\mathbf{a} \times \mathbf{c}} \otimes (\hat{I} - \hat{\sigma}_{\mathbf{c}})], \tag{D4b}$$

$$\frac{1}{2}\hat{\sigma}_{\mathbf{c}} \otimes \hat{I}(\hat{\sigma}_{\mathbf{a}} \otimes \hat{I})\hat{U}_{\mathbf{c}}^\dagger = \frac{1}{4}[\mathbf{a} \cdot \mathbf{c}\,(\hat{I} \otimes \hat{I} + \hat{I} \otimes \hat{\sigma}_{\mathbf{c}}$$
$$+ 2\hat{\sigma}_{\mathbf{c}} \otimes \hat{I} - 2\hat{\sigma}_{\mathbf{c}} \otimes \hat{\sigma}_{\mathbf{c}})$$
$$- \hat{\sigma}_{\mathbf{a}} \otimes (\hat{I} - \hat{\sigma}_{\mathbf{c}})$$
$$\times i\hat{\sigma}_{\mathbf{c} \times \mathbf{a}} \otimes (\hat{I} + \hat{\sigma}_{\mathbf{c}})], \tag{D4c}$$

$$-\frac{1}{2}\hat{\sigma}_{\mathbf{c}} \otimes \hat{\sigma}_{\mathbf{c}}(\hat{\sigma}_{\mathbf{a}} \otimes \hat{I})\hat{U}_{\mathbf{c}}^\dagger = \frac{1}{4}[\mathbf{a} \cdot \mathbf{c}\,(-\hat{I} \otimes \hat{I} - \hat{I} \otimes \hat{\sigma}_{\mathbf{c}}$$
$$+ 2\hat{\sigma}_{\mathbf{c}} \otimes \hat{I} - 2\hat{\sigma}_{\mathbf{c}} \otimes \hat{\sigma}_{\mathbf{c}})$$
$$- \hat{\sigma}_{\mathbf{a}} \otimes (\hat{I} - \hat{\sigma}_{\mathbf{c}})$$
$$- i\hat{\sigma}_{\mathbf{c} \times \mathbf{a}} \otimes (\hat{I} + \hat{\sigma}_{\mathbf{c}})]. \tag{D4d}$$

Adding (D4a)–(D4d) gives (D2).

Now consider $\hat{\rho}_{\mathrm{i}}^{(1)} = \hat{U}_{\mathrm{prep}}\hat{\rho}_0^{(1)}\hat{U}_{\mathrm{prep}}^\dagger$, where $\hat{\rho}_0^{(1)} = [\hat{\sigma}_{\mathbf{r}_0} \otimes \hat{I}^{\otimes(n-1)} + \cdots + \hat{I}^{\otimes(n-1)} \otimes \hat{\sigma}_{\mathbf{r}_0}]/N$. Since the preparatory unitary is symmetric under interchange pair of qubits it suffices to evaluate $\hat{U}_{\mathrm{prep}}\hat{\sigma}_{\mathbf{r}_0} \otimes \hat{I}^{\otimes(n-1)}\hat{U}_{\mathrm{prep}}^\dagger$. The $\hat{U}_{\mathbf{c}}$ factors in $\hat{U}_{\mathrm{prep}}$ and $\hat{U}_{\mathrm{prep}}^\dagger$ commute among each other and all of those that do not involve the leftmost qubit commute with the factors of the identity. Therefore, for each such pair of qubits a factor of $\hat{U}_{\mathbf{c}}$ multiplies a factor of $\hat{U}_{\mathbf{c}}^\dagger$, leaving the identity. Thus we need only consider the factors of $\hat{U}_{\mathbf{c}}$ in $\hat{U}_{\mathrm{prep}}$ that involve the leftmost qubit. This process can be illustrated with a three qubit system. Invoking (D2) gives that the $\hat{U}_{\mathbf{c}}$ acting on the leftmost and center qubits produces

$$\hat{\sigma}_{\mathbf{r}_0} \otimes \hat{I} \otimes \hat{I} \mapsto \hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{c}} \otimes \hat{I} + (\mathbf{r}_0 \cdot \mathbf{c})\hat{\sigma}_{\mathbf{c}} \otimes \hat{I} \otimes \hat{I}$$

TABLE I. Traces of square of terms from Eq. (37). The multiplicity indicates the number of times each type of terms occurs in the product $(\hat{\rho}_f^{(1)})^2$.

Term	Multiplicity	Contribution
1st	1	$\mathbf{r}_0^\top \dot{M}^\top \dot{M} \mathbf{r}_0 / N$
2nd or 3rd	$n-1$	$\mathbf{c}^\top \dot{M}^\top \dot{M} \mathbf{c} / N$
4th	1	$\mathbf{r}_0^\top P_{\mathbf{c}} \dot{M}^\top \dot{M} P_{\mathbf{c}} \mathbf{r}_0 / N$
5th	1	$n^2 \mathbf{r}_0^\top P_{\mathbf{c}} \dot{M}^\top \dot{M} P_{\mathbf{c}} \mathbf{r}_0 / N$

TABLE II. Traces of cross-term products from Eq. (37). The multiplicity indicates the number of times each type of terms occurs in the product $(\hat{\rho}_f^{(1)})^2$.

Term	Multiplicity	Contribution
1st and 2nd or 3rd	$2(n-1)$	$\mathbf{r}_0^\top P_{\mathbf{c}} \dot{M}^\top \dot{M} \mathbf{r}_0 / N$
1st and 4th	1	0
1st and 5th	2	$-n \, \mathbf{r}_0^\top P_{\mathbf{c}} \dot{M}^\top \dot{M} \mathbf{r}_0 / N$
2nd and 3rd	$(n-1)(n-2)$	$\mathbf{r}_0^\top P_{\mathbf{c}} \dot{M}^\top \dot{M} P_{\mathbf{c}} \mathbf{r}_0 / N$
2nd or 3rd and 4th	$n-1$	0
2nd or 3rd and 5th	$2(n-1)$	$-n \, \mathbf{r}_0^\top P_{\mathbf{c}} \dot{M}^\top \dot{M} P_{\mathbf{c}} \mathbf{r}_0 / N$
4th and 5th	1	0

$$-(\mathbf{r}_0 \cdot \mathbf{c}) \hat{\sigma}_{\mathbf{c}} \otimes \hat{\sigma}_{\mathbf{c}} \otimes \hat{I}. \tag{D5}$$

Now, noting that $\hat{U}_{\mathbf{c}}$ commutes with terms of the form $\hat{\sigma}_{\mathbf{c}} \otimes \hat{I}$ and $\hat{\sigma}_{\mathbf{c}} \otimes \hat{\sigma}_{\mathbf{c}}$, consider the factor $\hat{U}_{\mathbf{c}}$ acting on the leftmost and rightmost qubits. This leaves the second and third terms in Eq. (D5) unaltered since it commutes with them. However, according to (D2) it acts on the first term of (D5) to produce

$$\hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{c}} \otimes \hat{I} \mapsto \hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{c}} \otimes \hat{\sigma}_{\mathbf{c}} + (\mathbf{r}_0 \cdot \mathbf{c}) \hat{\sigma}_{\mathbf{c}} \otimes \hat{\sigma}_{\mathbf{c}} \otimes \hat{I}$$
$$-(\mathbf{r}_0 \cdot \mathbf{c}) \hat{\sigma}_{\mathbf{c}} \otimes \hat{\sigma}_{\mathbf{c}} \otimes \hat{\sigma}_{\mathbf{c}}. \tag{D6}$$

Thus substituting this into the first term of (D5) reveals that under all factors of $\hat{U}_{\mathbf{c}}$ the terms $\hat{\sigma}_{\mathbf{c}} \otimes \hat{\sigma}_{\mathbf{c}} \otimes \hat{I}$ cancel. Thus, under the preparatory unitary,

$$\hat{\sigma}_{\mathbf{r}_0} \otimes \hat{I} \otimes \hat{I} \mapsto \hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{c}} \otimes \hat{\sigma}_{\mathbf{c}} + (\mathbf{r}_0 \cdot \mathbf{c}) \hat{\sigma}_{\mathbf{c}} \otimes \hat{I} \otimes \hat{I}$$
$$-(\mathbf{r}_0 \cdot \mathbf{c}) \hat{\sigma}_{\mathbf{c}} \otimes \hat{\sigma}_{\mathbf{c}} \otimes \hat{\sigma}_{\mathbf{c}}. \tag{D7}$$

This pattern continues and, for arbitrary numbers of qubits,

$$\hat{U}_{\text{prep}} \left(\hat{\sigma}_{\mathbf{r}_0} \otimes \hat{I}^{\otimes(n-1)} \right) \hat{U}_{\text{prep}}^\dagger = \hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{c}}^{\otimes(n-1)}$$
$$+ (\mathbf{r}_0 \cdot \mathbf{c}) \, \hat{\sigma}_{\mathbf{c}} \otimes \hat{I}^{\otimes(n-1)}$$
$$- (\mathbf{r}_0 \cdot \mathbf{c}) \, \hat{\sigma}_{\mathbf{c}}^{\otimes n}. \tag{D8}$$

This determines the term $\hat{U}_{\text{prep}} \hat{\sigma}_{\mathbf{r}_0} \otimes \hat{I}^{\otimes(n-1)} \hat{U}_{\text{prep}}^\dagger$. This result, the symmetry of $\hat{U}_{\text{prep}}$ under interchange of qubits, the form of $\hat{\rho}_0^{(1)}$ and its symmetry under interchange of qubits, then give the first order channel input term of (35).

APPENDIX E: LOWEST ORDER SYMMETRIC PAIRWISE CORRELATED PROTOCOL QFI

To prove the result of Eq. (38), note that it emerges from $H = r^2 H^{(2)} + O(r^3)$ and the trace operation of Eq. (21), which requires computing the trace of the square of the entire right-hand side of Eq. (37). This results in a sum of the trace of each term squared together with the traces of all "cross terms"; we evaluate and list these separately. In both cases a useful tool is that $\text{Tr}\,[\hat{\sigma}_{\mathbf{a}} \hat{\sigma}_{\mathbf{b}}] = 2\, \mathbf{a}^\top \mathbf{b}$ for any vectors $\mathbf{a}$ and $\mathbf{b}$. Also note that the result will contain terms of the form $(\mathbf{r}_0 \cdot \mathbf{c}) \dot{M} \mathbf{c}$ and this can be expressed as $\dot{M} P_{\mathbf{c}} \, \mathbf{r}_0$ where $P_{\mathbf{c}}$ is the projection operator onto $\mathbf{c}$.

Then the squared terms of Eq. (37) are listed in Table I.

Similarly consider the "cross terms" of Eq. (37) are listed in Table II.

Adding these gives

$$H^{(2)} = N \, \text{Tr} \left[\left(\frac{\partial \hat{\rho}_f^{(1)}}{\partial \lambda} \right)^2 \right] = \mathbf{r}_0^\top [\dot{M}^\top \dot{M} + (3-n) P_{\mathbf{c}} \dot{M}^\top \dot{M} P_{\mathbf{c}}$$

$$-2 P_{\mathbf{c}} \dot{M}^\top \dot{M}] \mathbf{r}_0 + (n-1) \mathbf{c}^\top \dot{M}^\top \dot{M} \mathbf{c}. \tag{E1}$$

Note that $\mathbf{r}_0^\top P_{\mathbf{c}} \dot{M}^\top \dot{M} \mathbf{r}_0 = \mathbf{r}_0^\top \dot{M}^\top \dot{M} P_{\mathbf{c}} \mathbf{r}_0$ and thus a symmetric expression is

$$H^{(2)} = \mathbf{r}_0^\top [\dot{M}^\top \dot{M} + (3-n) P_{\mathbf{c}} \dot{M}^\top \dot{M} P_{\mathbf{c}}$$
$$- P_{\mathbf{c}} \dot{M}^\top \dot{M} - \dot{M}^\top \dot{M} P_{\mathbf{c}}] \mathbf{r}_0 + (n-1) \mathbf{c}^\top \dot{M}^\top \dot{M} \mathbf{c}. \tag{E2}$$

Algebra then gives the stated result.

APPENDIX F: BOUND ON THE LOWEST ORDER TERM IN THE QFI FOR UNITAL CHANNELS

The singular value decomposition is $\dot{M} = ASB$, with $S = \sum_i s_i P_i$ arranged so that $s_1 \geqslant s_2 \geqslant s_3 \geqslant 0$ and where P_i is a projector onto the unit vector $\hat{\mathbf{e}}_i$ and each of these is one of $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$. Then Eq. (38) becomes

$$H = r^2 \sum_i s_i^2 \big\{ \mathbf{r}_0^\top (I - P_{\mathbf{c}}) B^\top P_i B (I - P_{\mathbf{c}}) \mathbf{r}_0$$
$$+ (n-1) \mathbf{c}^\top P_i \mathbf{c} - (n-2) \mathbf{r}_0^\top P_{\mathbf{c}} B^\top P_i B P_{\mathbf{c}} \mathbf{r}_0 \big\}. \tag{F1}$$

The lower bound of Eq. (39) can be established by the particular choice of $\mathbf{c} = B^\top \hat{\mathbf{e}}_1$ and $\mathbf{r}_0 = B^\top \hat{\mathbf{e}}_2$. The upper bound can be established by noting that, since $s_1^2 \geqslant s_2^2 \geqslant s_3^2 \geqslant 0$,

$$H r^2 \leqslant s_1^2 \sum_i \big\{ \mathbf{r}_0^\top (I - P_{\mathbf{c}}) B^\top P_i B (I - P_{\mathbf{c}}) \mathbf{r}_0$$
$$+ (n-1) \mathbf{c}^\top B^\top P_i B \mathbf{c}$$
$$- (n-2) \mathbf{r}_0^\top P_{\mathbf{c}} B^\top P_i B P_{\mathbf{c}} \mathbf{r}_0 \big\}. \tag{F2}$$

Then the facts that $\sum_i P_i = I$, and $\mathbf{c}$, $\mathbf{r}_0$ are unit vectors and $P_{\mathbf{c}}$, $I - P_{\mathbf{c}}$ are projectors give

$$H \leqslant r^2 s_1^2 [n - (n-1) \mathbf{r}_0^\top P_{\mathbf{c}} \mathbf{r}_0]. \tag{F3}$$

The left side attains a maximum of $n s_1^2$ when $\mathbf{c}$ and $\mathbf{r}_0$ are perpendicular. This proves the result for the upper bound. It also implies that for the upper bound to be saturated the initial state Bloch-sphere vector direction and control direction must be perpendicular. But it does not guarantee that the upper bound can be saturated and, if not, it makes no statement about the directions of these vectors in order to attain the maximum QFI.

APPENDIX G: OPTIMAL QFI FOR $s_2 = s_3 = 0$

To prove the result that the optimal choice of control and initial-state directions is one where they are perpendicular, consider any fixed choice of $\mathbf{c}$. Then within Eq. (F1) there appears the term $W := \mathbf{r}_0^\top (I - P_{\mathbf{c}}) P_1 (I - P_{\mathbf{c}}) \mathbf{r}_0 + (n - 1)\mathbf{c}^\top P_1 \mathbf{c} - (n - 2)\mathbf{r}_0^\top P_{\mathbf{c}} P_1 P_{\mathbf{c}} \mathbf{r}_0$ and we will show that this is maximized when $\mathbf{r}_0$ and $\mathbf{c}$ are perpendicular. Note that $\mathbf{c}$ and $\hat{\mathbf{e}}_1$ span a plane. Then $\mathbf{r}_0$ can be decomposed into a vector perpendicular to the plane $\mathbf{r}_0^\perp$ and a vector parallel to the plane $\mathbf{r}_0^\parallel$. Neither of these necessarily has unit norm. Also let ϕ be the angle from $\hat{\mathbf{e}}_1$ to $\mathbf{c}$ and θ be the angle from $\mathbf{c}$ to $\mathbf{r}_0^\parallel$. As these three vectors lie in the same plane the angles can be chosen so that the from $\hat{\mathbf{e}}_1$ to $\mathbf{r}_0^\parallel$ is $\theta + \phi$. Then vector algebra gives

$$P_1 P_{\mathbf{c}} \mathbf{r}_0 = r_0^\parallel \cos\theta \cos\phi \, \hat{\mathbf{e}}_1 \hat{\mathbf{e}}_1, \tag{G1a}$$

$$P_1 (I - P_{\mathbf{c}}) \mathbf{r}_0 = -r_0^\parallel \sin\theta \sin\phi, \tag{G1b}$$

where $r_0^\parallel$ is the magnitude of $\mathbf{r}_0^\parallel$. Then

$$\mathbf{r}_0^\top P_{\mathbf{c}} P_1 P_{\mathbf{c}} \mathbf{r}_0 = (r_0^\parallel)^2 \cos^2\theta \cos^2\phi,$$

$$\mathbf{r}_0^\top (I - P_{\mathbf{c}}) P_1 (I - P_{\mathbf{c}}) \mathbf{r}_0 = (r_0^\parallel)^2 \sin^2\theta \sin^2\phi. \tag{G2a}$$

Separately $\mathbf{c}^\top P_1 \mathbf{c} = \cos^2\phi$. Thus

$$\begin{aligned} W &= (r_0^\parallel)^2 [\sin^2\theta \sin^2\phi + (2 - n)\cos^2\theta \cos^2\phi] \\ &\quad + (n - 1)\cos^2\phi \\ &= (r_0^\parallel)^2 \{\cos^2\theta [(3 - n)\cos^2\phi - 1] - \cos^2\phi\} \\ &\quad + (n - 1)\cos^2\phi. \end{aligned} \tag{G3}$$

For a given channel and choice of control direction $\mathbf{c}$, the variable ϕ is fixed and this must be optimized with respect to $\mathbf{r}_0$, i.e., with respect to θ and $r_0^\parallel$. Here, noting that for $n \geqslant 2$, the term $(3 - n)\cos^2\phi - 1 \leqslant 0$. This implies that the factor multiplying $(r_0^\parallel)^2$ is never positive. So the maximum for W_1 is attained when $r_0^\parallel = 0$. This gives $W_1 = (n - 1)\cos^2\phi$ and this attains a maximum of $n - 1$ when $\mathbf{c}$ is perpendicular to $\hat{\mathbf{e}}_1$.

APPENDIX H: HIGHER ORDER QFI TERMS FOR UNITAL CHANNELS

Equation (16) and that facts that $\hat{\rho}_{\mathrm{f}}^{(0)} = \hat{I}^{\otimes n}/N$ and $\hat{L}^{(0)} = 0$ imply that, to determine the third and fourth order terms in the QFI, we will need both $\hat{\rho}_{\mathrm{f}}^{(j)}$ and $\hat{L}^{(j)}$ for $j = 1, 2, 3$. Then Eqs. (15) and (20) and $\hat{\rho}_{\mathrm{f}}^{(0)} = \hat{I}^{\otimes n}/N$ give

$$\frac{\partial \hat{\rho}_{\mathrm{f}}^{(2)}}{\partial \lambda} = \frac{1}{N} \hat{L}^{(2)} + \frac{N}{2} \left\{ \frac{\partial \hat{\rho}_{\mathrm{f}}^{(1)}}{\partial \lambda}, \hat{\rho}_{\mathrm{f}}^{(1)} \right\}, \tag{H1}$$

where $\{,\}$ indicates the anticommutator. Thus

$$\hat{L}^{(2)} = N \frac{\partial \hat{\rho}_{\mathrm{f}}^{(2)}}{\partial \lambda} - \frac{N^2}{2} \frac{\partial}{\partial \lambda} \left[\hat{\rho}_{\mathrm{f}}^{(1)} \right]^2. \tag{H2}$$

Repeating this process gives

$$\frac{\partial \hat{\rho}_{\mathrm{f}}^{(3)}}{\partial \lambda} = \frac{1}{N} \hat{L}^{(3)} + \frac{N}{2} \frac{\partial}{\partial \lambda} \left\{ \hat{\rho}_{\mathrm{f}}^{(2)}, \hat{\rho}_{\mathrm{f}}^{(1)} \right\}$$

$$- \frac{N^2}{4} \left\{ \frac{\partial}{\partial \lambda} \left[\hat{\rho}_{\mathrm{f}}^{(1)} \right]^2, \hat{\rho}_{\mathrm{f}}^{(1)} \right\}. \tag{H3}$$

Thus

$$\hat{L}^{(3)} = N \frac{\partial \hat{\rho}_{\mathrm{f}}^{(3)}}{\partial \lambda} - \frac{N^2}{2} \frac{\partial}{\partial \lambda} \left\{ \hat{\rho}_{\mathrm{f}}^{(2)}, \hat{\rho}_{\mathrm{f}}^{(1)} \right\}$$

$$+ \frac{N^3}{4} \left\{ \frac{\partial}{\partial \lambda} \left[\hat{\rho}_{\mathrm{f}}^{(1)} \right]^2, \hat{\rho}_{\mathrm{f}}^{(1)} \right\}. \tag{H4}$$

Then Eqs. (16) and (20), the fact that $\hat{\rho}_{\mathrm{f}}^{(0)} = \hat{I}^{\otimes n}/N$ and the preceding expressions for the SLD terms give

$$\begin{aligned} H^{(3)} &= \mathrm{Tr}\left[\hat{L}^{(1)} \frac{\partial \hat{\rho}_{\mathrm{f}}^{(2)}}{\partial \lambda} + \hat{L}^{(2)} \frac{\partial \hat{\rho}_{\mathrm{f}}^{(1)}}{\partial \lambda} \right] \\ &= 2N \, \mathrm{Tr}\left[\frac{\partial \hat{\rho}_{\mathrm{f}}^{(1)}}{\partial \lambda} \frac{\partial \hat{\rho}_{\mathrm{f}}^{(2)}}{\partial \lambda} \right] \\ &\quad - \frac{N^2}{2} \mathrm{Tr}\left[\frac{\partial (\hat{\rho}_{\mathrm{f}}^{(1)})^2}{\partial \lambda} \frac{\partial \hat{\rho}_{\mathrm{f}}^{(1)}}{\partial \lambda} \right]. \end{aligned} \tag{H5}$$

Thus

$$H^{(3)} = 2N \, \mathrm{Tr}\left[\frac{\partial \hat{\rho}_{\mathrm{f}}^{(1)}}{\partial \lambda} \frac{\partial \hat{\rho}_{\mathrm{f}}^{(2)}}{\partial \lambda} \right] - N^2 \, \mathrm{Tr}\left[\frac{\partial (\hat{\rho}_{\mathrm{f}}^{(1)})^2}{\partial \lambda} \hat{\rho}_{\mathrm{f}}^{(1)} \right]. \tag{H6}$$

The fourth order term in the QFI can be obtained in a similar fashion, eventually giving

$$\begin{aligned} H^{(4)} &= N \, \mathrm{Tr}\left[2 \frac{\partial \hat{\rho}_{\mathrm{f}}^{(1)}}{\partial \lambda} \frac{\partial \hat{\rho}_{\mathrm{f}}^{(3)}}{\partial \lambda} + \left(\frac{\partial \hat{\rho}_{\mathrm{f}}^{(2)}}{\partial \lambda} \right)^2 \right] \\ &\quad - \frac{N^2}{2} \, \mathrm{Tr}\left[\frac{\partial \hat{\rho}_{\mathrm{f}}^{(1)}}{\partial \lambda} \frac{\partial}{\partial \lambda} \left\{ \hat{\rho}_{\mathrm{f}}^{(1)}, \hat{\rho}_{\mathrm{f}}^{(2)} \right\} + \frac{\partial \hat{\rho}_{\mathrm{f}}^{(2)}}{\partial \lambda} \frac{\partial (\hat{\rho}_{\mathrm{f}}^{(1)})^2}{\partial \lambda} \right] \\ &\quad + \frac{N^3}{4} \, \mathrm{Tr}\left[\frac{\partial \hat{\rho}_{\mathrm{f}}^{(1)}}{\partial \lambda} \left\{ \frac{\partial}{\partial \lambda} \left[\hat{\rho}_{\mathrm{f}}^{(1)} \right]^2, \hat{\rho}_{\mathrm{f}}^{(1)} \right\} \right]. \end{aligned} \tag{H7}$$

Further algebra yields

$$\begin{aligned} H^{(4)} &= N \, \mathrm{Tr}\left[2 \frac{\partial \hat{\rho}_{\mathrm{f}}^{(1)}}{\partial \lambda} \frac{\partial \hat{\rho}_{\mathrm{f}}^{(3)}}{\partial \lambda} + \left(\frac{\partial \hat{\rho}_{\mathrm{f}}^{(2)}}{\partial \lambda} \right)^2 \right] \\ &\quad - N^2 \, \mathrm{Tr}\left[\left(\frac{\partial \hat{\rho}_{\mathrm{f}}^{(1)}}{\partial \lambda} \right)^2 \hat{\rho}_{\mathrm{f}}^{(2)} + \frac{\partial \hat{\rho}_{\mathrm{f}}^{(2)}}{\partial \lambda} \frac{\partial (\hat{\rho}_{\mathrm{f}}^{(1)})^2}{\partial \lambda} \right] \\ &\quad + \frac{N^3}{4} \, \mathrm{Tr}\left[\left(\frac{\partial (\hat{\rho}_{\mathrm{f}}^{(1)})^2}{\partial \lambda} \right)^2 \right]. \end{aligned} \tag{H8}$$

The second and third order terms in the channel input state can be determined by repeatedly using Eq. (D2) and the additional result

$$\begin{aligned} \hat{U}_{\mathbf{c}}(\hat{\sigma}_{\mathbf{a}} \otimes \hat{\sigma}_{\mathbf{b}}) \hat{U}_{\mathbf{c}}^\dagger &= (\mathbf{a} \times \mathbf{c}) \cdot \hat{\boldsymbol{\sigma}} \otimes (\mathbf{b} \times \mathbf{c}) \cdot \hat{\boldsymbol{\sigma}} \\ &\quad + (\mathbf{a} \cdot \mathbf{c}) \hat{I} \otimes \hat{\sigma}_{\mathbf{b}} + (\mathbf{b} \cdot \mathbf{c}) \hat{\sigma}_{\mathbf{a}} \otimes \hat{I} \\ &\quad + (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{c})(\hat{\sigma}_{\mathbf{c}} \otimes \hat{\sigma}_{\mathbf{c}} \end{aligned}$$

$$-\hat{\sigma}_{\mathbf{c}} \otimes \hat{I} - \hat{I} \otimes \hat{\sigma}_{\mathbf{c}}). \tag{H9}$$

To prove this, note that $\hat{U}_{\mathbf{c}}(\hat{\sigma}_{\mathbf{a}} \otimes \hat{\sigma}_{\mathbf{b}})\hat{U}_{\mathbf{c}}^{\dagger} = \hat{U}_{\mathbf{c}}(\hat{\sigma}_{\mathbf{a}} \otimes \hat{I})\hat{U}_{\mathbf{c}}\hat{U}_{\mathbf{c}}^{\dagger}(\hat{I} \otimes \hat{\sigma}_{\mathbf{b}})\hat{U}_{\mathbf{c}}^{\dagger}$ since $\hat{U}_{\mathbf{c}}\hat{U}_{\mathbf{c}}^{\dagger} = \hat{I}$. Then invoking (D2) twice and using (D1) repeatedly gives (H9).

We apply these to determine the second and third order terms in the channel input density operator provided that *the initial state Bloch-sphere direction and the control direction are perpendicular*. Equation (35) implies that

$$\hat{\rho}_{\mathrm{i}}^{(1)} = \frac{1}{N}\left(\hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{c}}^{\otimes(n-1)} + \cdots + \hat{\sigma}_{\mathbf{c}}^{\otimes(n-1)} \otimes \hat{\sigma}_{\mathbf{r}_0}\right) \tag{H10}$$

and this contains all terms with a single factor of $\hat{\sigma}_{\mathbf{r}_0}$ and $n-1$ factors of $\hat{\sigma}_{\mathbf{c}}$.

The second order term in the initial state is

$$\hat{\rho}_0^{(2)} = \frac{1}{N}\left[\hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{r}_0} \otimes \hat{I}^{\otimes(n-2)} + \cdots + \hat{I}^{\otimes(n-2)} \otimes \hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{r}_0}\right], \tag{H11}$$

where every possible permutation including exactly two factors of $\hat{\sigma}_{\mathbf{r}_0}$ appears with the brackets. Then consider the effect of the preparatory unitary on $\hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{r}_0} \otimes \hat{I}^{\otimes(n-2)}$. We label the qubits from left to right as $1, 2, 3, \ldots, n$. Then $\hat{U}_{\mathbf{c}}$ acting on any pair of qubits in the range $3, 4, \ldots, n$ has no effect on this term. Now consider $\hat{U}_{\mathbf{c}}$ acting on qubits 1 and 3. Since $\mathbf{r}_0$ is perpendicular to $\mathbf{c}$, Eq. (D2) shows that this produces $\hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{c}} \otimes \hat{I}^{\otimes(n-3)}$. Now consider the subsequent action of $\hat{U}_{\mathbf{c}}$ on qubits 2 and 3. According to Eq. (H9) this results in $\hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{r}_0} \otimes \hat{I}^{\otimes(n-2)}$. Thus the only factor of $\hat{U}_{\mathbf{c}}$ that will have a nontrivial effect on this terms is that acting on qubits 1 and 2. Equation (H9) shows that this gives $\hat{\sigma}_{\mathbf{r}_0 \times \mathbf{c}} \otimes \hat{\sigma}_{\mathbf{r}_0 \times \mathbf{c}} \otimes \hat{I}^{\otimes(n-2)}$. Thus, if $\mathbf{r}_0$ is perpendicular to $\mathbf{c}$, the second order channel input term is

$$\hat{\rho}_{\mathrm{i}}^{(2)} = \frac{1}{N}\left[\hat{\sigma}_{\mathbf{r}_0 \times \mathbf{c}} \otimes \hat{\sigma}_{\mathbf{r}_0 \times \mathbf{c}} \otimes \hat{I}^{\otimes(n-2)}\right.$$
$$\left. + \cdots + \hat{I}^{\otimes(n-2)} \otimes \hat{\sigma}_{\mathbf{r}_0 \times \mathbf{c}} \otimes \hat{\sigma}_{\mathbf{r}_0 \times \mathbf{c}}\right], \tag{H12}$$

which contains terms with every possible arrangement of two factors of $\hat{\sigma}_{\mathbf{r}_0 \times \mathbf{c}}$.

The third order initial state term is

$$\hat{\rho}_0^{(3)} = \frac{1}{N}\left[\hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{r}_0} \otimes \hat{I}^{\otimes(n-3)}\right.$$
$$\left. + \cdots + \hat{I}^{\otimes(n-3)} \otimes \hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{r}_0}\right] \tag{H13}$$

and this contains terms with every possible arrangement of three factors of $\hat{\sigma}_{\mathbf{r}_0}$. The effects of the preparatory unitary on this can be determined by considering $\hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{r}_0} \otimes \hat{I}^{\otimes(n-3)}$. Then by the argument for the second order term the factors of $\hat{U}_{\mathbf{c}}$ acting on all pairs of qubits from 2 to n produce $\hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{r}_0 \times \mathbf{c}} \otimes \hat{\sigma}_{\mathbf{r}_0 \times \mathbf{c}} \otimes \hat{I}^{\otimes(n-3)}$. It remains to consider the effect of each $\hat{U}_{\mathbf{c}}$ which involves qubit 1. By Eq. (D2) the effect of $\hat{U}_{\mathbf{c}}$ for qubits 1 and 4, $\hat{U}_{\mathbf{c}}$ for qubits 1 and 5 up to $\hat{U}_{\mathbf{c}}$ for qubits 1 and n is to produce $\hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{r}_0 \times \mathbf{c}} \otimes \hat{\sigma}_{\mathbf{r}_0 \times \mathbf{c}} \otimes \hat{\sigma}_{\mathbf{c}}^{\otimes(n-3)}$. Now the effect of $\hat{U}_{\mathbf{c}}$ for qubits 1 and 3 on this is to produce $\hat{\sigma}_{\mathbf{r}_0 \times \mathbf{c}} \otimes \hat{\sigma}_{\mathbf{r}_0 \times \mathbf{c}} \otimes \hat{\sigma}_{(\mathbf{r}_0 \times \mathbf{c}) \times \mathbf{c}} \otimes \hat{\sigma}_{\mathbf{c}}^{\otimes(n-3)}$ which is the same as $-\hat{\sigma}_{\mathbf{r}_0 \times \mathbf{c}} \otimes \hat{\sigma}_{\mathbf{r}_0 \times \mathbf{c}} \otimes \hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{c}}^{\otimes(n-3)}$. Finally, the effect of $\hat{U}_{\mathbf{c}}$ for qubits 1 and 2 on this is to produce $-\hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{c}}^{\otimes(n-3)}$. Thus we obtain the third order term in the channel

input state,

$$\hat{\rho}_{\mathrm{i}}^{(3)} = -\frac{1}{N}\left[\hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{c}}^{\otimes(n-3)}\right.$$
$$\left. + \cdots + \hat{\sigma}_{\mathbf{c}}^{\otimes(n-3)} \otimes \hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{r}_0}\right], \tag{H14}$$

where this contains all terms with three factors of $\hat{\sigma}_{\mathbf{r}_0}$.

We now consider the situation where the channel acts on once on the leftmost qubit and aim to compute the third and fourth orders terms in the QFI via Eqs. (H6) and (H8). The key tools are that $\mathrm{Tr}\,[\hat{\sigma}_{\mathbf{a}}] = 0$, $\mathrm{Tr}\,[\hat{\sigma}_{\mathbf{a}}\hat{\sigma}_{\mathbf{b}}] = 2\mathbf{a} \cdot \mathbf{b}$ and that the trace of a product state is the product of the partial traces of the factors. These immediately imply that all terms in $H^{(3)}$ and $H^{(4)}$ that contain a factor of $\hat{\rho}_{\mathrm{f}}^{(2)}$ or $\frac{\partial \hat{\rho}_{\mathrm{f}}^{(2)}}{\partial \lambda}$ multiplied by factors that contain $\hat{\rho}_{\mathrm{f}}^{(1)}$ or its derivative trace to zero, since these always contain factors of $\mathrm{Tr}\,[\hat{\sigma}_{\mathbf{r}_0 \times \mathbf{c}}]$, $\mathrm{Tr}\,[\hat{\sigma}_{\mathbf{r}_0 \times \mathbf{c}}\hat{\sigma}_{\mathbf{r}_0}]$, or $\mathrm{Tr}\,[\hat{\sigma}_{\mathbf{r}_0 \times \mathbf{c}}\hat{\sigma}_{\mathbf{c}}]$. Similarly the product $\frac{\partial \hat{\rho}_{\mathrm{f}}^{(1)}}{\partial \lambda}\frac{\partial \hat{\rho}_{\mathrm{f}}^{(3)}}{\partial \lambda}$ contains at least one factor of $\hat{\sigma}_{\mathbf{r}_0}$ and this also traces to zero. Thus, in this scenario,

$$H^{(3)} = -N^2\,\mathrm{Tr}\left[\frac{\partial\left(\hat{\rho}_{\mathrm{f}}^{(1)}\right)^2}{\partial \lambda}\hat{\rho}_{\mathrm{f}}^{(1)}\right], \tag{H15a}$$

$$H^{(4)} = N\,\mathrm{Tr}\left[\left(\frac{\partial \hat{\rho}_{\mathrm{f}}^{(2)}}{\partial \lambda}\right)^2\right]$$
$$+ \frac{N^3}{4}\,\mathrm{Tr}\left[\left(\frac{\partial\left(\hat{\rho}_{\mathrm{f}}^{(1)}\right)^2}{\partial \lambda}\right)^2\right]. \tag{H15b}$$

Now consider $\frac{\partial(\hat{\rho}_{\mathrm{f}}^{(1)})^2}{\partial \lambda}\hat{\rho}_{\mathrm{f}}^{(1)}$. Within the rightmost $n-1$ factors there will always be at least one factor of $\hat{\sigma}_{\mathbf{c}}$ or $\hat{\sigma}_{\mathbf{r}_0}$. Thus this traces to zero and Eq. (H15a) gives that if $\mathbf{r}_0$ and $\mathbf{c}$ are perpendicular then $H^{(3)} = 0$.

In order to compute $H^{(4)}$, we first consider $\left(\frac{\partial \hat{\rho}_{\mathrm{f}}^{(2)}}{\partial \lambda}\right)^2$. Here

$$\frac{\partial \hat{\rho}_{\mathrm{f}}^{(2)}}{\partial \lambda} = \frac{1}{N}\left[\dot{M}(\mathbf{r}_0 \times \mathbf{c}) \cdot \hat{\boldsymbol{\sigma}} \otimes \hat{\sigma}_{\mathbf{r}_0 \times \mathbf{c}} \otimes \hat{I}^{\otimes(n-2)}\right.$$
$$+ \dot{M}(\mathbf{r}_0 \times \mathbf{c}) \cdot \hat{\boldsymbol{\sigma}} \otimes \hat{I} \otimes \hat{\sigma}_{\mathbf{r}_0 \times \mathbf{c}} \otimes \hat{I}^{\otimes(n-3)}$$
$$\left. + \cdots + \dot{M}(\mathbf{r}_0 \times \mathbf{c}) \cdot \hat{\boldsymbol{\sigma}} \otimes \hat{I}^{\otimes(n-2)} \otimes \hat{\sigma}_{\mathbf{r}_0 \times \mathbf{c}}\right].$$

After squaring, the cross terms from this each contain at least one factor of $\hat{\sigma}_{\mathbf{r}_0 \times \mathbf{c}}$ and do not contribute to the trace. The $n-1$ squared terms are all $(\mathbf{r}_0 \times \mathbf{c})^{\top}\dot{M}^{\top}\dot{M}(\mathbf{r}_0 \times \mathbf{c})\hat{I}^{\otimes n}/N^2$. Thus

$$\mathrm{Tr}\left[\left(\frac{\partial \hat{\rho}_{\mathrm{f}}^{(2)}}{\partial \lambda}\right)^2\right] = \frac{n-1}{N}(\mathbf{r}_0 \times \mathbf{c})^{\top}\dot{M}^{\top}\dot{M}(\mathbf{r}_0 \times \mathbf{c}). \tag{H16}$$

The remaining term involves

$$\hat{\rho}_{\mathrm{f}}^{(1)} = \frac{1}{N}\left[(M\mathbf{r}_0) \cdot \hat{\boldsymbol{\sigma}} \otimes \hat{\sigma}_{\mathbf{c}}^{\otimes(n-1)}\right.$$
$$+ (M\mathbf{c}) \cdot \hat{\boldsymbol{\sigma}} \otimes \hat{\sigma}_{\mathbf{r}_0} \otimes \hat{\sigma}_{\mathbf{c}}^{\otimes(n-2)}$$
$$\left. + \cdots + (M\mathbf{c}) \cdot \hat{\boldsymbol{\sigma}} \otimes \hat{\sigma}_{\mathbf{c}}^{\otimes(n-2)} \otimes \hat{\sigma}_{\mathbf{r}_0}\right]. \tag{H17}$$

Thus for $n \geqslant 3$,

$$\left(\hat{\rho}_{\mathrm{f}}^{(1)}\right)^2 = \frac{1}{N^2}\Big[\mathbf{r}_0^\top M^\top M\mathbf{r}_0\,\hat{I}^{\otimes n} + (n-1)\mathbf{c}^\top M^\top M\mathbf{c}\,\hat{I}^{\otimes n} + (n-1)\mathbf{c}^\top M^\top M\mathbf{c}\left(\hat{I}\otimes\hat{\sigma}_{\mathbf{r}_0\times\mathbf{c}}\otimes\hat{\sigma}_{\mathbf{r}_0\times\mathbf{c}}\otimes\hat{I}^{\otimes(n-3)}\right.$$
$$+\cdots\hat{I}^{\otimes(n-2)}\otimes\hat{\sigma}_{\mathbf{r}_0\times\mathbf{c}}\otimes\hat{\sigma}_{\mathbf{r}_0\times\mathbf{c}}\big) + (M\mathbf{r}_0\times M\mathbf{c})\cdot\hat{\boldsymbol{\sigma}}\otimes\hat{\sigma}_{\mathbf{r}_0\times\mathbf{c}}\otimes\hat{I}^{\otimes(n-2)}$$
$$+\cdots+(M\mathbf{r}_0\times M\mathbf{c})\cdot\hat{\boldsymbol{\sigma}}\otimes\hat{I}^{\otimes(n-2)}\otimes\hat{\sigma}_{\mathbf{r}_0\times\mathbf{c}}\Big]. \tag{H18}$$

After differentiating and squaring, only the squares contribute to the trace. Thus

$$\mathrm{Tr}\left[\left(\frac{\partial\left(\hat{\rho}_{\mathrm{f}}^{(1)}\right)^2}{\partial\lambda}\right)^2\right] = \frac{1}{N^3}\left[\mathbf{r}_0^\top\frac{\partial(M^\top M)}{\partial\lambda}\mathbf{r}_0 + (n-1)\,\mathbf{c}^\top\frac{\partial(M^\top M)}{\partial\lambda}\mathbf{c}\right]^2$$
$$+\frac{4(n-1)}{N^3}\left|\frac{\partial}{\partial\lambda}(M\mathbf{r}_0\times M\mathbf{c})\right|^2 + \frac{2(n-1)(n-2)}{N^3}\left[\mathbf{c}^\top\frac{\partial(M^\top M)}{\partial\lambda}\mathbf{c}\right]^2, \tag{H19}$$

where $|\mathbf{a}|^2 := \mathbf{a}\cdot\mathbf{a}$. Thus

$$H^{(4)} = (n-1)\,(\mathbf{r}_0\times\mathbf{c})^\top\dot{M}^\top\dot{M}(\mathbf{r}_0\times\mathbf{c})\frac{1}{4}\left[\mathbf{r}_0^\top\frac{\partial(M^\top M)}{\partial\lambda}\mathbf{r}_0 + (n-1)\,\mathbf{c}^\top\frac{\partial(M^\top M)}{\partial\lambda}\mathbf{c}\right]^2$$
$$+ (n-1)\left|\frac{\partial}{\partial\lambda}(M\mathbf{r}_0\times M\mathbf{c})\right|^2 + \frac{(n-1)(n-2)}{2}\left[\mathbf{c}^\top\frac{\partial(M^\top M)}{\partial\lambda}\mathbf{c}\right]^2. \tag{H20}$$

This scales as n^2. For example, for the depolarizing channel, $M = \lambda I$ and thus

$$H^{(4)} = (n-1) + \lambda^2 n(3n-2). \tag{H21}$$

APPENDIX I: MEASUREMENTS THAT SATURATE THE QFI

Suppose that the single channel invocation is followed by application of $\hat{U}_{\mathrm{prep}}^\dagger = \hat{U}_{\mathrm{prep}}$ and, after this each qubit is measured in a particular basis. We will consider the situation where $\mathbf{r}_0$ and $\mathbf{c}$ are perpendicular and show that for a particular set of single-qubit measurements, the resulting classical Fisher information gives the bounds of Eq. (39). For $\mathbf{r}_0$ and $\mathbf{c}$ perpendicular, prior to channel invocation the state of the system is, to first order in the purity,

$$\hat{\rho}_{\mathrm{i}} = \hat{\rho}_{\mathrm{i}}^{(0)} + r\hat{\rho}_{\mathrm{i}}^{(1)}, \tag{I1}$$

where $\hat{\rho}_{\mathrm{i}}^{(0)} = \hat{I}^{\otimes n}/N$ and

$$\hat{\rho}_{\mathrm{i}}^{(1)} = \frac{1}{N}\left(\hat{\sigma}_{\mathbf{r}_0}\otimes\hat{\sigma}_{\mathbf{c}}^{\otimes(n-1)} + \cdots + \hat{\sigma}_{\mathbf{c}}^{\otimes(n-1)}\otimes\hat{\sigma}_{\mathbf{r}_0}\right). \tag{I2}$$

Here the sum contains every possible term with only one factor of $\hat{\sigma}_{\mathbf{r}_0}$.

Under a unital channel, $\hat{I}$ remains invariant and thus a single channel invocation on the leftmost qubit results in the state

$$\hat{\rho}_{\mathrm{f}} = \hat{\rho}_{\mathrm{f}}^{(0)} + r\hat{\rho}_{\mathrm{f}}^{(1)}, \tag{I3}$$

where $\hat{\rho}_{\mathrm{f}}^{(0)} = \hat{I}^{\otimes n}/N$ and

$$\hat{\rho}_{\mathrm{f}}^{(1)} = \frac{1}{N}\big(\hat{\sigma}_{\mathbf{a}}\otimes\hat{\sigma}_{\mathbf{c}}^{\otimes(n-1)} + \hat{\sigma}_{\mathbf{b}}\otimes\hat{\sigma}_{\mathbf{r}_0}\otimes\hat{\sigma}_{\mathbf{c}}^{\otimes(n-2)}$$
$$+\cdots+\hat{\sigma}_{\mathbf{b}}\otimes\hat{\sigma}_{\mathbf{c}}^{\otimes(n-2)}\otimes\hat{\sigma}_{\mathbf{r}_0}\big). \tag{I4}$$

Here, for convenience, $\mathbf{a} := M\mathbf{r}_0$ and $\mathbf{b} := M\mathbf{c}$.

The premeasurement state of the system is then $\hat{\rho}_{\mathrm{m}} := \hat{U}_{\mathrm{prep}}^\dagger\hat{\rho}_{\mathrm{f}}\hat{U}_{\mathrm{prep}}$ and

$$\hat{\rho}_{\mathrm{m}} = \hat{\rho}_{\mathrm{m}}^{(0)} + r\hat{\rho}_{\mathrm{m}}^{(1)}, \tag{I5}$$

where $\hat{\rho}_{\mathrm{m}}^{(i)} := \hat{U}_{\mathrm{prep}}^\dagger\hat{\rho}_{\mathrm{f}}^{(i)}\hat{U}_{\mathrm{prep}}$. Then

$$\hat{\rho}_{\mathrm{m}}^{(0)} = \frac{1}{N}\,\hat{I}^{\otimes n}, \tag{I6}$$

since $\hat{U}_{\mathrm{prep}}$ is unitary.

In order to determine $\hat{\rho}_{\mathrm{m}}^{(1)}$, note that $\hat{U}_{\mathrm{prep}}$ is a product of one $\hat{U}_{\mathbf{c}}$ for each pair of qubits. We will determine $\hat{\rho}_{\mathrm{m}}^{(1)}$ in two steps: (a) evaluate the effect of every factor of $\hat{U}_{\mathbf{c}}$ that does not include the leftmost qubit and (b) evaluate the effect of every factor of $\hat{U}_{\mathbf{c}}$ that does contain the leftmost qubit.

First, consider the effect of every factor of $\hat{U}_{\mathbf{c}}$ that does not include the leftmost qubit. We need to consider two types of terms in $\hat{\rho}_{\mathrm{f}}^{(1)}$. One type is $\hat{\sigma}_{\mathbf{a}}\otimes\hat{\sigma}_{\mathbf{c}}^{\otimes(n-1)}$ and this stays invariant. The other type contains a single factor of $\hat{\sigma}_{\mathbf{r}_0}$ in the rightmost $n-1$ factors of the tensor product. As an example consider $\hat{\sigma}_{\mathbf{b}}\otimes\hat{\sigma}_{\mathbf{r}_0}\otimes\hat{\sigma}_{\mathbf{c}}^{\otimes(n-2)}$. Determining the effect of the factors of $\hat{U}_{\mathbf{c}}$ requires only determining the effect of the factors that involve the second qubit from the left. Then using Eq. (H9) and the fact that $\mathbf{r}_0$ and $\mathbf{c}$ are perpendicular gives

$$\hat{U}_{\mathbf{c}}\hat{\sigma}_{\mathbf{r}_0}\otimes\hat{\sigma}_{\mathbf{c}}\hat{U}_{\mathbf{c}}^\dagger = \hat{\sigma}_{\mathbf{r}_0}\otimes\hat{I}. \tag{I7}$$

Repeatedly using this gives that every factor of $\hat{U}_{\mathbf{c}}$ that does not include the leftmost qubit maps

$$\hat{\sigma}_{\mathbf{b}}\otimes\hat{\sigma}_{\mathbf{r}_0}\otimes\hat{\sigma}_{\mathbf{c}}^{\otimes(n-2)} \mapsto \hat{\sigma}_{\mathbf{b}}\otimes\hat{\sigma}_{\mathbf{r}_0}\otimes\hat{I}^{\otimes(n-2)}. \tag{I8}$$

Thus every factor of $\hat{U}_{\mathbf{c}}$ that does not include the leftmost qubit maps

$$\hat{\rho}_{\mathrm{f}}^{(1)} \mapsto \frac{1}{N}\big(\hat{\sigma}_{\mathbf{a}}\otimes\hat{\sigma}_{\mathbf{c}}^{\otimes(n-1)} + \hat{\sigma}_{\mathbf{b}}\otimes\hat{\sigma}_{\mathbf{r}_0}\otimes\hat{I}^{\otimes(n-2)}$$
$$+\cdots+\hat{\sigma}_{\mathbf{b}}\otimes\hat{I}^{\otimes(n-2)}\otimes\hat{\sigma}_{\mathbf{r}_0}\big). \tag{I9}$$

Second, consider the remaining factors of $\hat{U}_{\mathbf{c}}$ that each include the leftmost qubit. In order to assess these, we use Eq. (H9) to give

$$\hat{U}_{\mathbf{c}}\hat{\sigma}_{\mathbf{a}} \otimes \hat{\sigma}_{\mathbf{c}}\hat{U}_{\mathbf{c}}^{\dagger} = \hat{\sigma}_{\mathbf{a}} \otimes \hat{I}$$
$$+ (\mathbf{a} \cdot \mathbf{c})(\hat{\sigma}_{\mathbf{c}} \otimes \hat{\sigma}_{\mathbf{c}} - \hat{\sigma}_{\mathbf{c}} \otimes \hat{I}), \quad \text{(I10)}$$

$$\hat{U}_{\mathbf{c}}\hat{\sigma}_{\mathbf{b}} \otimes \hat{\sigma}_{\mathbf{r}_0}\hat{U}_{\mathbf{c}}^{\dagger} = (\mathbf{b} \times \mathbf{c}) \cdot \hat{\boldsymbol{\sigma}} \otimes (\mathbf{r}_0 \times \mathbf{c}) \cdot \hat{\boldsymbol{\sigma}}$$
$$+ (\mathbf{b} \cdot \mathbf{c})\hat{I} \otimes \hat{\sigma}_{\mathbf{r}_0}, \quad \text{(I11)}$$

$$\hat{U}_{\mathbf{c}}\hat{\sigma}_{\mathbf{b}} \otimes \hat{I}\hat{U}_{\mathbf{c}}^{\dagger} = \hat{\sigma}_{\mathbf{b}} \otimes \hat{\sigma}_{\mathbf{c}}$$
$$+ (\mathbf{b} \cdot \mathbf{c})(\hat{\sigma}_{\mathbf{c}} \otimes \hat{I} - \hat{\sigma}_{\mathbf{c}} \otimes \hat{\sigma}_{\mathbf{c}}). \quad \text{(I12)}$$

Repeatedly applying these gives that under $\hat{U}_{\mathbf{c}}$ acting on every pair that includes the leftmost qubit,

$$\hat{\sigma}_{\mathbf{a}} \otimes \hat{\sigma}_{\mathbf{c}}^{\otimes(n-1)} \mapsto \hat{\sigma}_{\mathbf{a}} \otimes \hat{I}^{\otimes(n-1)} + (\mathbf{a} \cdot \mathbf{c})\hat{\sigma}_{\mathbf{c}}^{\otimes n}$$
$$- (\mathbf{a} \cdot \mathbf{c})\hat{\sigma}_{\mathbf{c}} \otimes \hat{I}^{\otimes(n-1)}. \quad \text{(I13)}$$

The remaining terms can be evaluated by considering $\hat{\sigma}_{\mathbf{b}} \otimes \hat{\sigma}_{\mathbf{r}_0} \otimes \hat{I}^{\otimes(n-2)}$. Again applying the results of Eq. (H9) gives

$$\hat{\sigma}_{\mathbf{b}} \otimes \hat{\sigma}_{\mathbf{r}_0} \otimes \hat{I}^{\otimes(n-2)} \mapsto (\mathbf{b} \times \mathbf{c}) \cdot \hat{\boldsymbol{\sigma}} \otimes (\mathbf{r}_0 \times \mathbf{c}) \cdot \hat{\boldsymbol{\sigma}} \otimes \hat{\sigma}_{\mathbf{c}}^{\otimes(n-2)}$$
$$\times (\mathbf{b} \cdot \mathbf{c})\hat{I} \otimes \hat{\sigma}_{\mathbf{r}_0} \otimes \hat{I}^{\otimes(n-2)}. \quad \text{(I14)}$$

Extending these results to all the permutations in Eq. (I9) then gives

$$\hat{\rho}_{\mathrm{m}}^{(1)} = \frac{1}{N}\big[\hat{\sigma}_{\mathbf{a}} \otimes \hat{I}^{\otimes(n-1)}(\mathbf{a} \cdot \mathbf{c})\big(\hat{\sigma}_{\mathbf{c}}^{\otimes n} - \hat{\sigma}_{\mathbf{c}} \otimes \hat{I}^{\otimes(n-1)}\big)$$
$$\times (\mathbf{b} \times \mathbf{c}) \cdot \hat{\boldsymbol{\sigma}} \otimes (\mathbf{r}_0 \times \mathbf{c}) \cdot \hat{\boldsymbol{\sigma}} \otimes \hat{\sigma}_{\mathbf{c}}^{\otimes(n-2)}$$
$$+ \cdots + (\mathbf{b} \times \mathbf{c}) \cdot \hat{\boldsymbol{\sigma}} \otimes \hat{\sigma}_{\mathbf{c}}^{\otimes(n-2)} \otimes (\mathbf{r}_0 \times \mathbf{c}) \cdot \hat{\boldsymbol{\sigma}}$$
$$\times (\mathbf{b} \cdot \mathbf{c})\big(\hat{I} \otimes \hat{\sigma}_{\mathbf{r}_0} \otimes \hat{I}^{\otimes(n-2)}\big)$$
$$+ \cdots + \hat{I}^{\otimes(n-1)} \otimes \hat{\sigma}_{\mathbf{r}_0}\big)\big]. \quad \text{(I15)}$$

Now consider measurement for each qubit along the Bloch-sphere direction $\mathbf{r}_0$. The corresponding single-qubit projection operators are

$$\hat{\Pi}_{\pm} := \tfrac{1}{2}\big(\hat{I} \pm \hat{\sigma}_{\mathbf{r}_0}\big) \quad \text{(I16)}$$

and we denote the associated measurement outcomes by $+$ or $-$. Since $\hat{\rho}_{\mathrm{m}}^{(1)}$ is invariant under interchange between any of the rightmost $n - 1$, the probabilities of the measurement outcomes depend on: (a) whether the outcome for the leftmost qubit is $+$ or $-$ and (b) the number k of $+$ outcomes amongst the rightmost $n - 1$ qubits. To this end let $p(+, k)$ be the probability that the measurement outcome for the leftmost qubit is $+$ and k of the outcomes amongst the rightmost $n - 1$ qubits are $+$. Similarly let $p(-, k)$ be the probability that the

measurement outcome for the leftmost qubit is $-$ and k of the outcomes among the rightmost $n - 1$ qubits are $+$.

A simplifying aspect of computing these probabilities is the fact that if $\mathbf{r}_0$ and $\mathbf{c}$ are perpendicular, then $\mathrm{Tr}\,(\hat{\Pi}_{\pm}\hat{\sigma}_{\mathbf{c}}) = 0$. Thus only the first and last series of terms on the right-hand side of Eq. (I15) contribute to the probabilities. Now consider $p(+, k)$. One way of attaining this outcome is a $+$ for the leftmost qubit, followed by $+$ outcomes for the next k qubits and finally $-$ outcomes for the rightmost $n - 1 - k$ qubits. The probability with which this occurs is

$$\frac{1}{2^n}\,\mathrm{Tr}\,\big[\big(\hat{I} + \hat{\sigma}_{\mathbf{r}_0}\big) \otimes \cdots \otimes \big(\hat{I} + \hat{\sigma}_{\mathbf{r}_0}\big)$$
$$\otimes\big(\hat{I} - \hat{\sigma}_{\mathbf{r}_0}\big) \otimes \cdots \otimes \big(\hat{I} - \hat{\sigma}_{\mathbf{r}_0}\big)\hat{\rho}_{\mathrm{m}}\big], \quad \text{(I17)}$$

where there are $k + 1$ factors with the $+$ sign and $n - k - 1$ with the $-$ sign. Using Eqs. (I6) and (I15) then gives that this returns

$$\frac{1}{N}\,[1 + r(\mathbf{r}_0 \cdot \mathbf{a}) + r(\mathbf{r}_0 \cdot \mathbf{c})(2k - n + 1)]. \quad \text{(I18)}$$

This, the fact that there are $\binom{n-1}{k}$ ways to attain k outcomes of $+$ among the rightmost $n - 1$ qubits and the definitions of $\mathbf{a}$ and $\mathbf{b}$, give

$$p(+, k) = \frac{1}{N}\binom{n-1}{k}[1 + r\mathbf{r}_0^{\top} M \mathbf{r}_0$$
$$+ r\mathbf{c}^{\top} M \mathbf{c}(2k - n + 1)]. \quad \text{(I19)}$$

Similarly

$$p(-, k) = \frac{1}{N}\binom{n-1}{k}[1 - r\mathbf{r}_0^{\top} M \mathbf{r}_0$$
$$+ r\mathbf{c}^{\top} M \mathbf{c}(2k - n + 1)]. \quad \text{(I20)}$$

The classical Fisher information in this case is

$$F(\lambda) = \sum_{k=0}^{n-1}\left[\frac{1}{p(+, k)}\left(\frac{\partial p(+, k)}{\partial \lambda}\right)^2\right.$$
$$\left.+ \frac{1}{p(-, k)}\left(\frac{\partial p(-, k)}{\partial \lambda}\right)^2\right]. \quad \text{(I21)}$$

Substituting and retaining only the lowest order nonzero terms in the purity gives

$$F(\lambda) = r^2[(\mathbf{r}_0^{\top}\dot{M}\mathbf{r}_0)^2 + (n - 1)(\mathbf{c}^{\top}\dot{M}\mathbf{c})^2]. \quad \text{(I22)}$$

Now consider the choices for $\mathbf{r}_0$ and $\mathbf{c}$ that resulted in the bounds of Eq. (39). Following the notation of Appendix F, choosing $\mathbf{c} = B^{\top}\hat{\mathbf{e}}_1$ and $\mathbf{r}_0 = B^{\top}\hat{\mathbf{e}}_2$, gives

$$F(\lambda) = r^2(n - 1)s_1^2 + r^2 s_2^2. \quad \text{(I23)}$$

Thus this particular set of Bloch-sphere direction and measurement choices saturates the lower bound of the QFI.

[1] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).

[2] C. M. Caves, Phys. Rev. D **23**, 1693 (1981).

[3] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. **72**, 3439 (1994).

[4] M. Sarovar and G. J. Milburn, J. Phys. A **39**, 8487 (2006).

[5] V. Giovannetti, S. Lloyd, and L. Maccone, Phys. Rev. Lett. **96**, 010401 (2006).

[6] M. Paris, Int. J. Quantum. Inform. **7**, 125 (2009).

[7] P. Kok and B. W. Lovett, *Introduction to Optical Quantum Information Processing* (Cambridge University Press, Cambridge, 2010).

[8] V. Giovannetti, S. Lloyd, and L. Maccone, Nat. Photon. **5**, 222 (2011).

[9] J. Kołodyński and R. Demkowicz-Dobrzański, New J. Phys. **15**, 073043 (2013).

[10] A. Fujiwara, Phys. Rev. A **63**, 042304 (2001).

[11] M. Sasaki, M. Ban, and S. M. Barnett, Phys. Rev. A **66**, 022308 (2002).

[12] M. Frey, D. Collins, and K. Gerlach, J. Phys. A **44**, 205306 (2011).

[13] A. Fujiwara and H. Imai, J. Phys. A **36**, 8093 (2003).

[14] Z. Ji, G. Wang, R. Duan, Y. Feng, and M. Ying, IEEE Trans. Inf. Theory **54**, 5172 (2008).

[15] A. Fujiwara, Phys. Rev. A **70**, 012317 (2004).

[16] G. M. D'Ariano, C. Macchiavello, and P. Perinotti, Phys. Rev. A **72**, 042327 (2005).

[17] A. Datta, L. Zhang, N. Thomas-Peter, U. Dorner, B. J. Smith, and I. A. Walmsley, Phys. Rev. A **83**, 063836 (2011).

[18] K. Modi, H. Cable, M. Williamson, and V. Vedral, Phys. Rev. X **1**, 021022 (2011).

[19] O. Pinel, P. Jian, N. Treps, C. Fabre, and D. Braun, Phys. Rev. A **88**, 040102 (2013).

[20] D. Collins, Phys. Rev. A **87**, 032301 (2013).

[21] D. Collins and J. Stephens, Phys. Rev. A **92**, 032324 (2015).

[22] H. Cable, M. Gu, and K. Modi, Phys. Rev. A **93**, 040304 (2016).

[23] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

[24] H. Cramér, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, NJ, 1946).

[25] C. O'Loan, arXiv:1001.3971.

[26] H. Nagaoka, in *Asymptotic Theory of Quantum Statistical Inference*, edited by M. Hayashi (World Scientific, Singapore, 2012), Chap. 9, pp. 113–124.

[27] Y. Watanabe, in *Quantum Estimation Theory*, Springer Theses (Springer, Japan, 2014), pp. 37–44.

[28] F. Toscano, W. P. Bastos, and R. L. de Matos Filho, Phys. Rev. A **95**, 042125 (2017).

[29] O. E. Barndorff-Nielsen and R. D. Gill, J. Phys. A **33**, 4481 (2000).

[30] M. A. Ballester, Phys. Rev. A **69**, 022303 (2004).

[31] M. Hotta, T. Karasawa, and M. Ozawa, Phys. Rev. A **72**, 052334 (2005).

[32] B. M. Escher, R. L. de Matos Filho, and L. Davidovich, Nat. Phys. **7**, 406 (2011).

[33] M. Zwierz, C. A. Pérez-Delgado, and P. Kok, Phys. Rev. A **85**, 042112 (2012).

[34] R. Demkowicz-Dobrzański and L. Maccone, Phys. Rev. Lett. **113**, 250801 (2014).

[35] W. Dür, M. Skotiniotis, F. Fröwis, and B. Kraus, Phys. Rev. Lett. **112**, 080801 (2014).

[36] Z. Huang, C. Macchiavello, and L. Maccone, Phys. Rev. A **94**, 012101 (2016).

[37] P. Cappellaro, J. Emerson, N. Boulant, C. Ramanathan, S. Lloyd, and D. G. Cory, Phys. Rev. Lett. **94**, 020502 (2005).

[38] D. Cory, R. Laflamme, E. Knill, L. Viola, T. Havel, N. Boulant, G. Boutis, E. Fortunato, S. Lloyd, R. Martinez, C. Negrevergne, M. Pravia, Y. Sharf, G. Teklemariam, Y. Weinstein, and W. Zurek, Fortschr. Phys. **48**, 875 (2000).

[39] C. Ramanathan, N. Boulant, Z. Chen, D. G. Cory, I. Chuang, and M. Steffen, Quant. Info. Proc. **3**, 15 (2004).

[40] J. A. Jones, Prog. NMR Spectrosc. **59**, 91 (2011).

[41] S. F. Huelga, C. Macchiavello, T. Pellizzari, A. K. Ekert, M. B. Plenio, and J. I. Cirac, Phys. Rev. Lett. **79**, 3865 (1997).

[42] M. Jarzyna and R. Demkowicz-Dobrzański, New J. Phys. **17**, 013010 (2015).